

**UNIVERSITETET I OSLO**

**Institutt for informatikk**

**Semiotiske systemlyder**

**Masteroppgave**

(60 studiepoeng)

Bjørnar Hjelmevold

**1. mai 2009**



# Sammendrag

---

Denne oppgaven studerer den semiotiske betydningen som ligger i systemlydene til Microsoft Windows. Lyder fra tre utgaver av Windows blir analysert: Windows 2000, Windows XP og Windows Vista. Formålet med analysen er å se hvordan betydningen har utviklet seg fra et system til et annet, og om lyddesignet er styrt av konvensjoner eller basert på kreative tilfeldigheter. I et bredere perspektiv forsøker oppgaven å besvare hvordan betydningen i lyd forhandles fram til brukeren opp mot samtidige inntrykk som skrift og bilde, og det drøftes hvilken rolle systemlyder kan spille for å gjøre et brukergrensesnitt mer forståelig. Det teoretiske utgangspunktet for oppgaven ligger i at vi oppfatter lyd som objekter i bevegelse, og at bevegelsesinntrykkene kan knyttes opp mot en bestemt handling som har skjedd i brukergrensesnittet. Fra dette bevegelsesinntrykket kan det utledes en betydning i systemlyden, men denne betydningen eksisterer også multimodalt opp mot andre samtidige inntrykk.

## Abstract

---

In this thesis, system sounds from three versions of Microsoft Windows are analyzed semiotically in order to determine their signified meaning. The analysis includes system sounds from Windows 2000, Windows XP, and Windows Vista. The purpose of the analysis is to see if the sound design is bound by conventions or just based on creative coincidences. In a broader perspective, this thesis tries to answer how meaning is negotiated to the user in relation to juxtaposed text and image. The findings are then subject to a discussion about what role system sounds can play in order to make user interfaces easier to understand. The theoretical basis is the premise that we perceive sound as objects in motion, and that these images of movement can be connected to a specific action in the user interface. Not only do these images of movement allow signified meaning in the system sounds, but meaning can also be signified multimodally in relation to other juxtaposed modes of expression.

# Innholdsfortegnelse

---

|   |                  |
|---|------------------|
| <b><u>1. Innledning</u></b>                     | <b><u>1</u></b>  |
| 1.1 Problemstilling                             | 3                |
| 1.2 Oppgavens struktur                          | 4                |
| <b><u>2. Hvordan analysere systemlyder?</u></b> | <b><u>6</u></b>  |
| 2.1 Lytteperspektiv                             | 6                |
| 2.1.1 Redusert lytting                          | 7                |
| 2.2 Handling-lyd                                | 7                |
| 2.2.1 Lydobjektet                               | 8                |
| 2.2.2 Den lydskapende bevegelsen                | 9                |
| 2.2.3 Lydobjektets typologi og morfologi        | 10               |
| 2.3 Multimodal semiotikk                        | 11               |
| 2.3.1 Det semiotiske tegnet                     | 11               |
| 2.3.2 Tegnet i annen potens                     | 12               |
| 2.3.3 Forankring og avløsning                   | 13               |
| 2.3.4 Multimodalitet                            | 16               |
| 2.4 Metafor teori                               | 18               |
| <b><u>3. Deskriptiv semiotisk analyse</u></b>   | <b><u>22</u></b> |
| 3.1 Tekniske valg og kategorisering             | 22               |
| 3.1.1 Analyseverktøy                            | 22               |
| 3.1.2 Valg av lyder og deres kategorisering     | 24               |
| 3.1.3 Vurdering av lydhendelser                 | 26               |
| 3.1.4 Viktige unntak                            | 28               |
| 3.2 Deskriptiv semiotisk analyse                | 29               |
| 3.2.1 Oppsummerende tabell                      | 49               |
| 3.3 Problemstilling 1: Historisk utvikling      | 50               |

|  |                  |
|--|------------------|
| <b><u>4. Multimodal analyse</u></b>                  | <b><u>55</u></b> |
| 4.1 Sannhetsforsterkning                             | 55               |
| 4.2 Multimodal analyse                               | 57               |
| 4.3 Problemstilling 2: Systemlydenes betydningsstrøm | 67               |
| <b><u>5. Kontekstualisering</u></b>                  | <b><u>71</u></b> |
| 5.1 Microsofts systemlydformål                       | 71               |
| 5.2 Microsofts faremomenter                          | 74               |
| 5.2.1 Tekniske og miljømessige hensyn                | 74               |
| 5.2.2 Tilrettelegging for funksjonshemmede           | 76               |
| 5.2.3 Smak og kultur                                 | 76               |
| 5.3 Et forskersyn                                    | 77               |
| 5.3.1 Visjoner om lyd i brukergrensesnitt            | 78               |
| 5.3.2 SonicFinder                                    | 80               |
| 5.3.3 Sammenligning av Gaver og Microsoft            | 81               |
| 5.4 Kløften mellom forskning og programvareutvikling | 82               |
| <b><u>6. Avslutning</u></b>                          | <b><u>85</u></b> |
| 6.1 Oppsummering                                     | 85               |
| 6.2 Veien videre                                     | 87               |
| <b><u>Referanser</u></b>                             | <b><u>89</u></b> |

# 1. Innledning

---

Tenk deg en helt vanlig dag. Du begynner dagen med lyden av vekkerklokken som ringer. Mobiltelefonen slås på, og den piper for å melde at du har fått en SMS-melding. Du ringer opp vedkommende som sendte deg meldingen, og midt i samtalen piper telefonen igjen fordi batterinivået begynner å bli lavt. Samtidig er maten ferdig varmet opp, og dette erfarer du når mikrobølgeovnen kimer frokosten inn. På vei til jobb står du på perrongen og venter på toget ditt, og du hører en tone fra høyttaleranlegget. Denne tonen gjør at du forventer å høre en vennlig stemme erklære at toget ditt er forsinket, men heldigvis var det bare en beskjed om at toget kommer inn på en annen plattform enn vanlig. Mens du venter, besøker du nærmeste kiosk for å kjøpe en avis. Selgeren bak disken bruker en strekkodeleser som piper når avisen er klar til salg, og du taster inn koden din på bankterminalen i en serie av korte pipelyder. Livet vårt er fullt av små og korte lyder som vi hører fra elektroniske apparater. Noen av disse lydene kjenner vi igjen og forstår, mens andre har en uforståelig betydning inntil vi får vite hva som har skjedd gjennom andre midler. Hvilken slags betydning kan ligge gjemt i disse lydene?

“Beeps can’t have specific meanings assigned to them” ([Microsoft 2008:708](#))

Sitatet ovenfor er noe tatt ut av konteksten<sup>1</sup>, men det setter ord på en oppfatning i samfunnet vårt om at en overflod av elektroniske lyder som høres like ut, kan gjøre det vanskelig for oss å definere en betydning i lydene. Forvirringen som da kan oppstå tenker jeg å illustrere ved å forlenge hverdageksemplet nevnt i forrige avsnitt. Vel framme på jobb setter du deg ned foran datamaskinen og prøver å sjekke e-posten din, men det eneste du opplever er at datamaskinen spiller en liten lyd for deg. Du forstår ikke av denne lyden hvorvidt maskinen har problemer med å hente inn ny e-post, om det ikke er noe ulest e-post å hente inn, eller om maskinen rett og slett ikke forstår hva du mener når du trykker tastekombinasjonen ‘Control-M’. Likevel vet du at *noe* har skjedd, fordi lyden fungerer som en varsellyd som tilkaller din oppmerksomhet. Denne lyden skiller seg fra den vanlige viftestøyen som kommer fra

---

<sup>1</sup> Microsoft skiller mellom pipelyder og spesialdesignede systemlyder slik som de som analyseres i denne oppgaven. Selskapet sier at pipelyder ikke kan bære en spesiell betydning, i motsetning til de lydene som Microsoft har laget spesielt for Windows Vista. Dette hevdes midt i en kontekst hvor det argumenteres for hvor viktig det er for programvareutviklere å benytte seg av Microsofts egne Vista-lyder. ([Microsoft 2008:708](#))

datamaskinen og fra andre lyder i omgivelsene, fordi den inntreffer samtidig som det vises en skriftlig melding på dataskjermen. Lyden oppleves ikke som om den stammer fra omgivelsene eller fra fysiske prosesser i maskinens interne komponenter, men heller som om den er en del av *brukergrensesnittet* som skal forenkle kommunikasjonen mellom deg og datamaskinen din. Dette brukergrensesnittet er en virtuell verden bestående av symboler som skrivebord, dokumenter, mapper, vinduer og andre metaforer som til sammen prøver å knytte datamaskinens operativsystem til begreper som vi kjenner fra vår virkelige verden. Lyden oppfattes som en del av denne virtuelle verdenen og regnes dermed som en *systemlyd* ([Microsoft 2008:703](#)).

Et brukergrensesnitt kan formidle mening til databrukeren via ulike midler, for eksempel via skrift, bilder og grafiske ikoner. I denne kombinasjonen av tegn ligger det mening som kan tolkes av brukeren, og *semiotikken* er læren om hvordan mening formidles fra slike tegn. Denne oppgaven bygger på en forutsetning om at systemlyder i likhet med skrift og bilder kan fungere som tegn, og at systemlydene dermed kan analyseres semiotisk. Imidlertid er det usikkert i hvilken grad en vanlig databruker ser lydens rolle som en meningsbærer, særlig hvis denne databrukeren daglig befinner seg i en kaotisk og forvirrende elektronisk lydverden, slik som i eksempelet innledningsvis i dette kapittelet. Utfordringen blir derfor å påvise betydning, selv i systemlyder som ved første oppfattelse ikke synes å ha noen betydning.

Det er ikke bare den vanlige databruker som kan undervurdere lyd som en meningsbærer. Innenfor forskning har det lenge eksistert en holdning om at det ikke går an å finne semiotisk betydning i lyd på samme måte som med skrift eller bilde, og at lydens egenskaper ikke er definert tilstrekkelig i semiotisk sammenheng ([Barthes 1977c](#)). Dette synes å ha ført til at lyd ikke vies like mye oppmerksomhet som designmessig virkemiddel i forhold til sine visuelle slektninger. Et eksempel på dette finner vi når en omfattende lærebok i interaksjonsdesign ([Sharp et al. 2007](#)) inneholder mye informasjon om hvordan brukergrensesnitt kan utformes visuelt, og hvordan betydning kan komme fram via ulike visuelle virkemidler. I samme lærebok begrenses informasjonen om lydbruk til et avsnitt hvor det forklares at lyd er et fint virkemiddel til å påvirke brukerens humør ([ibid.:](#)183). I den senere tid ser vi spor av at denne trenden er i ferd med å snu. Forskningstradisjoner som musikkteori, musikkpsykologi, nevrovitenskap, kognitiv teori og interaksjonsdesign har i løpet av de siste årtiene har økt sitt fokus på hvordan vi forstår lydfenomener ([Rocchesso og Fontana 2003:V](#)), og en tilsvarende

økende fokusering på lyd ser vi også innenfor semiotikken ([Maasø 1995](#), [Hull og Nelson 2005](#)). Vi ser nå tegn på at lyd begynner å bli tatt på alvor i forhold til sine historisk sett semiotiske storesøsken, skrift og bilde.

## 1.1 Problemstilling

I denne oppgaven skal jeg studere den semiotiske betydningen i et utvalg av systemlyder fra Microsoft Windows. Jeg har valgt dette analyseobjektet fordi det er det operativsystemet som med god margin har størst markedsandel i dagens PCer, og nå som flere og flere samfunnstjenester blir digitalisert, blir de fleste av oss databrukere, enten vi ønsker det eller ikke. Derfor er det sannsynlig at noen av systemlydene som jeg analyserer fra Microsoft Windows vil kunne gjenkjennes av leseren. Dette gjør det lettere for ham eller henne både å forstå hvilken betydning som kan ligge i systemlydene ut fra egne erfaringer. Det blir også lettere for leseren å overføre funnene til andre analoge situasjoner, for eksempel systemlyder i andre datamaskiner eller bærbare enheter som mobiltelefoner. Jeg har valgt å dele oppgavens problemstilling i to:

1. Hvordan har semiotiske trekk ved systemlyder i Microsoft Windows utviklet seg de siste årene?
2. Hva kjennetegner systemlydenes betydningsstrøm?

Den første problemstillingen har et historisk perspektiv, men det er ikke hensikten å etablere en historisk oversikt over systemlydenes betydning. Hensikten med å analysere historisk er både å se på om det finnes noen trender innenfor lyddesign, og å undersøke om systemlydenes betydning er såpass etablert at den er gitt gjennom konvensjoner. Et visuelt semiotisk eksempel på dette finner vi i betydningen til en trekant hvor et av hjørnene 'peker til høyre'. Dette tegnet kan ha flere betydninger, men er såpass abstrakt at vi forstår det gjennom et sett med kulturelle regler, nemlig konvensjoner. Hvis trekanten er plassert på en fjernkontrollknapp, kan tegnet bety 'play'. Er trekanten derimot plassert på et nødutgangskilt, betyr tegnet at skiltet gjelder for retningen 'høyre'. Disse to betydningene er vi nødt til å forstå gjennom konvensjoner som vi enten er blitt fortalt eller har funnet ut av på egen hånd. Jeg finner det spennende å undersøke om det finnes tilsvarende konvensjoner som gjelder for hvordan vi oppfatter betydning i systemlyder. For å finne ut av dette, kommer jeg til å analysere et utvalg

av systemlyder fra de tre nyeste versjonene av Windows: Windows 2000 ([Microsoft 1999](#)), Windows XP ([Microsoft 2001](#)) og Windows Vista ([Microsoft 2006](#)).

Gjennom den andre problemstillingen ønsker jeg å studere hvordan betydningen i en systemlyd formidles til lytteren. I den virkelige verden er vi vant med at lyd kommer fra en bestemt lydkilde som det på en eller annen måte har skjedd noe med. I brukergrensesnittet er det derimot langt mer komplisert, fordi handlingene og lydkildene der er konstruerte, og de har ikke nødvendigvis noen konkret referanse i den virkelige verden. Den interne betydningen i en systemlyd havner derfor i en kontekst som kan påvirke hvordan betydningen oppfattes hos brukeren. Det finnes også retninger innenfor moderne semiotikk som hevder at lydens betydning formes via dens relasjon til andre samtidige uttrykksformer som for eksempel skrift og bilde ([Kress og Van Leeuwen 2001](#)), og forskere som benytter seg av denne teorien har etterlyst en bedre forståelse av hvordan lyden passer inn i en slik sammenheng ([Hull og Nelson 2005:252](#)). Min andre problemstilling ser derfor ikke så mye på systemlydenes konkrete betydning, men fokuserer heller på lydens rolle som en meningsbærer, og hvordan tradisjonelle semiotiske syn på betydningsstrømmer mellom synkronstilt skrift og bilde utfordres når lyd inkluderes i denne sammenhengen.

## 1.2 Oppgavens struktur

Siden dette er en masteroppgave i digitale medier, kommer den i all hovedsak til å ha en medievitenskapelig innfallsvinkel: semiotikken og dens yngre variant sosialsemiotikken. Problemstillingene og oppgavestrukturen bærer preg av denne vinklingen, men det er noen viktige nyanser som bør nevnes. Tidligere i dette kapittelet nevnte jeg hvordan eldre semiotisk forskning har konsentrert seg om skrift og bilde, og at det først er i senere tid at lyd for alvor er blitt gjenstand for interesse. Denne utviklingen kan skyldes en trend i dagens forskning, hvor det er viktig å se en sak fra flere sider. Skal man kunne se en sak fra flere sider, kan det lønne seg å kombinere ulike forskningstradisjoner. Lydforståelse i brukergrensesnitt mener jeg er et såpass sammensatt emne at det kan være fruktbart å inkludere teori fra andre nærliggende fagtradisjoner som lydteori og interaksjonsdesign.

Jeg vil først etablere et teorigrunnlag på tvers av flere fagtradisjoner, dernest argumentere for hvordan de ulike teoriene kan kombineres og komme fram til en teorisammensetning som er fornuftig å bruke i analysen. Etter dette vil analyseobjektene bli studert på detaljnivå via en



deskriptiv analyse som har til hensikt å besvare den første problemstillingen om lydenes interne betydning. Neste utfordring blir å gjøre en analyse som skal besvare den andre problemstillingen, hvor jeg ser på hvordan betydningen i en systemlyd kommer fram i relasjon til andre samtidige inntrykk i brukergrensesnittet. Begge disse problemstillingene handler om lydens betydning, men det kan være interessant å se på de bakenforliggende årsakene til at det kan være slik. Derfor kommer det siste kapittelet før oppsummeringen til å kontekstualisere funnene mine, og kapittelet vil prøve å gi et bidrag til hvordan lyder kan designes i framtiden. Oppgavens fokus går derfor fra et bredt teorigrunnlag og ned til en smal verktøykasse med analysetermer, for så gradvis å bevege analyseperspektivet innenfra og ut ved først å se på intern betydning, til betydning opp mot andre synkrone inntrykk og til slutt betydning i forhold til analyseområdets kontekst.

Til slutt i dette innledende kapittelet vil jeg gjøre rede for min motivasjon bak å skrive denne oppgaven. Jeg er en masterstudent som har tidligere erfaring fra lyddesign, og min interesse for detaljer i enkeltlyder ligger nok derfor over gjennomsnittet for den jevne databruker. Med mitt utgangspunkt kan det virke som premissene er lagt for at jeg automatisk vil konkludere med at 'lyddesign er det som teller'. Noe slikt fornekter jeg ikke, og det er nettopp derfor jeg har valgt å skrive denne oppgaven. Det store spørsmålet som gjenstår, er likevel *hvordan* det går an å hevde at lyddesign er viktig å utvikle videre, noe denne oppgaven kommer til å ta for seg.

## 2. Hvordan analysere systemlyder

---

Systemlydenes semiotiske betydning er et komplisert felt å analysere, fordi det involverer så mange ulike forskningstradisjoner. Medievitenskap, musikkvitenskap og interaksjonsdesign har alle litt å bidra med, og som vi kommer til å se, vil jeg dessuten anvende noe metafor-teori som et bindeledd mellom de tre andre fagtradisjonene. Med så mange elementer kan det fort hende at det blir en del overlappende termer som forstås ulikt i hver av fagtradisjonene. I dette kapitlet skal jeg derfor se nærmere på en del slike termer, og stake ut noen fellesnevner som jeg kan bruke som verktøy i analysen min. Underveis vil jeg reflektere rundt metodebruk og argumentere for at metodene jeg har valgt er anvendbare på systemlyder i brukergrensesnitt.

### 2.1 Lytteperspektiv

Som forsker støter man lett på situasjoner hvor det å angripe et problem fra en annen vinkel kan gi helt andre resultater. Tatt i betraktning alle de ulike måtene vi mennesker nyter musikk på (melodi, rytme, tekst og så videre), kommer det fram hvor utslagsgivende det blir å velge et representativt lytteperspektiv.<sup>2</sup> Her ligger det et grunnleggende problem: ved normal bruk av datamaskiner sitter ikke brukeren og lytter nøye etter detaljer i lydene – jeg vil tro at de fleste av oss konsentrerer seg om det som blir vist på skjermen og hva vi ønsker å få gjort med det, og at lydinformasjonen kommer i andre rekke. Jeg som skal analysere lydene, har derimot allerede fra starten av inntatt et annet lytteperspektiv, hvor lydenes innhold og betydning kommer i *første* rekke. Se mitt lytteperspektiv i forhold til dets rake motsetning, nemlig den personen som konsekvent har skrudd datamaskinens lydvolume helt ned og følgelig aldri hører noen systemlyder i det hele tatt. Derfor har jeg per definisjon et utgangspunkt som jeg ikke kan argumentere for at er representativt overfor en typisk databruker, men derimot overfor en fiktiv databruker i en ideell verden hvor all betydning formidles etter planen.

---

<sup>2</sup> I tillegg til lyttetessig og analytisk perspektivforskjell, spiller naturligvis også tekniske og fysiske forskjeller som avspillingsutstyr og plassering i rom en viktig rolle.

### 2.1.1 Redusert lytting

Det lytteperspektivet som jeg kommer til å benytte meg av i den detaljerte analysen, heter *reduert lytting*. Opphavsmannen bak dette begrepet er Pierre Schaeffer, som definerer begrepet slik: se bort fra lydens opprinnelige kontekst, og fokuser i stedet på lydens soniske og rå bestanddeler (Schaeffer 1966:271). Schaeffers teori er fra 1960-tallet, en tid hvor strukturalismen dominerte europeiske (særlig franske og britiske) kulturstudier (Storey 2006). Schaeffer følger strukturalistenes tankegang når han velger å konsentrere seg om lydens enkle bestanddeler, og se bort fra hensyn til kompleksiteten i menneskelig lydoppfattelse, kontekstens betydning og informasjon gitt gjennom konvensjoner. Et motiv bak å ville studere lyd på nettopp denne måten kan ligge i de teknologiske nyvinningene som skjedde på 1900-tallet rundt lydopptak, lydsyntese og høyttalerteknikk. Schaeffers reduserte lytting er en teknikk som tar utgangspunkt i en lyttesituasjon hvor lydens opprinnelige kilder ikke er synlige for lytteren. Det samme kan sies om lyttesituasjonen når man opplever systemlyder spilt av fra høyttalerne til en datamaskin, hvor systemet som produserer lydene ikke er synlig for lytteren, og lydkilden står i et langt mer abstrakt forhold til den lyden som høres. Jeg mener derfor at det er relevant å benytte seg av en redusert lytting ikke bare for musikkanalyse, men også til å analysere systemlyder.

## 2.2 Handling-lyd

Den grunnleggende teorien fra musikkvitenskapen som jeg velger å benytte i denne oppgaven, tar utgangspunkt i økologisk persepsjonsforskning og anvender den i lydmessig sammenheng. Sentralt i økologisk persepsjonsforskning ligger tanken om at vi forstår verden ut fra vår kunnskap om hva ulike gjenstander kan brukes til. James J. Gibson (1979) lanserte begrepet *affordances* som forklarer at vi oppfatter verden rundt oss som et resultat av hva gjenstander tillater at vi gjør med dem. For eksempel kan en ball trilles, kastes og sittes på. En seng kan også sittes på, men den kan hverken trilles eller kastes, derimot kan man legge seg i den. Hvilke funksjoner som hvert objekt kan ha, forstår vi gjennom vår daglige erfaring med verden rundt oss, og hvilke *handlinger* vi gjør med dem. Tilsvarende erfaring kan komme godt med når vi skal forstå de virtuelle objektene i et brukergrensesnitt, og en modifisert<sup>3</sup> utgave av Gibsons *affordance*-begrep blir også brukt innenfor interaksjonsdesign (Norman [1988] 2002, Svanæs 2000).

---

<sup>3</sup> 'Affordance' i Normans interaksjonsdesign baserer seg mer på hva det er *sannsynlig* å bruke gjenstander til, enn på hva det er *mulig* å gjøre (Svanæs 2000:35-36)

Selv om Gibsons økologiske syn var myntet på visuell persepsjon, har koblingen mellom persepsjon og handling også funnet veien til musikkvitenskapen. [Alexander Refsum Jensenius \(2008\)](#) lanserer begrepet *handling-lyd* som en forbindelse vi oppfatter ut fra vår erfaring med hvordan lyder oppstår:

“The human capacity of imagining and predicting sounds seems to be based on our knowledge of the acoustical features of objects and the mechanical properties of actions” ([Jensenius 2008:23](#)).

Jensenius skiller mellom to typer handling-lyd: *handling-lyd-koblinger* og *handling-lyd-slektskap* ([ibid.:](#)21). Den første kategorien er direkte koblinger som beskriver hvordan lyd oppstår i naturen, mens *handling-lyd-slektskap* beskriver mer eller mindre troverdige kombinasjoner av handling og lyd, slike som oppstår i virtuelle verdener som for eksempel et brukergrensesnitt. I denne oppgaven vil mitt musikkteoretiske utgangspunkt være handling-lyd-slektskapet. Jeg nevnte i innledningen at datamaskiner kan lage lyder i brukergrensesnittet, men også maskinens mekaniske komponenter kan lage lyd (viftestøy), og ikke minst er det mye lyd i omgivelsene rundt datamaskinen. Siden handling-lyd-slektskapet tar utgangspunkt i økologisk persepsjon, vil lytteropplevelsen til brukeren også være farget av omgivelses- og maskinlyd. Derfor er det et poeng for meg å understreke at jeg i denne oppgaven ser bort fra lyder fra omgivelsene og fra selve datamaskinen, og at jeg kun fokuserer på lydene som inntreffer i brukergrensesnittets virtuelle verden. Bak handling-lyd-slektskapet til Jensenius ligger det en forståelse av at vi hører lyd som objekter<sup>4</sup>, og at objektene skaper lyd gjennom bevegelse. Lyd som objekter og lyd som bevegelse er fenomener som jeg nå vil se litt nærmere på.

### 2.2.1 Lydobjektet

Opp gjennom årene har forskere stilt seg kritiske til de skylappene som følger med å bruke redusert lytting som metode ([Chion \[1983\] 1995:30](#), [Nattiez 1990:95](#)), men det finnes også forskning som har kommet fram til at Schaeffers lytteprosess kan være hensiktsmessig ([Godøy 2006:150](#)), spesielt når man drar inn et annet av Schaeffers begreper, nemlig *lydobjektet*. Et lydobjekt kan forstås som alle perseptuelle lydfenomener og lydhendelser oppfattet som en enhet, hvor denne er et resultat av en mental prosess forårsaket av en redusert lytting

---

<sup>4</sup> Egentlig et objekt-handling-objekt-system, inspirert av Gibson ([Jensenius 2008:22](#))

(Schaeffer 1966:263, Godøy 2006:150). Enklere sagt er det lydobjekter man lytter etter når man foretar en redusert lytting.

En god grunn til å bruke lydobjekter i lyd- og musikkanalyse er at man får redusert et komplekst lydbilde ned til å bli en sammensetning av mindre komplekse byggesteiner. Jeg kommer imidlertid ikke til å benytte meg av denne muligheten i musikkanalyse, men heller analysere systemlyder som er såpass korte at de lett kan tolkes som et enkelt lydobjekt. Selv om musikken vi hører ved på- og avlogging av Microsoft Windows også er en systemlyd, er den mer kompleks i forhold til de andre korte systemlydene, og jeg mener at det å analysere lengre systemlyder vil kunne fortrenge små detaljer som jeg ellers ville funnet i de korte lydene. I brukergrensesnittet til en datamaskin forekommer systemlyder som et brudd på stillheten<sup>5</sup>, og er sånn sett selvstendige enheter med klare avgrensninger. Derfor vil jeg ikke gå nærmere inn på diskusjonen om segregering<sup>6</sup> av lydobjekter.

### 2.2.2 Den lydskapende bevegelsen

Det som gjør lydobjekter særlig interessante i en analyse av systemlyder, er måten et lydobjekt kan knyttes opp mot lydskapende handlinger som inntreffer i brukergrensesnittet. Det brobyggende begrepet som vi trenger å kjenne til her, er *den lydskapende gesten*. Med ‘gest’ menes i denne sammenheng en forestilt lydskapende bevegelse som vi danner oss mentalt ut fra lydobjektets kvaliteter (Godøy 2006). Gester brukes imidlertid innen veldig mange ulike fagtradisjoner, og det kan derfor være praktisk å forenkle termen til å være et *bevegelsesinntrykk* (Jensenius 2008:35). Det er også denne forenklete tolkningen som jeg kommer til å anvende videre i oppgaven.

Fra handling-lyd-forbindelsen nevnt tidligere ligger det en forståelse av at lyder som vi hører kan spores tilbake til en lydskapende handling, og at denne handlingen kan oppleves som en bevegelse. Ved å kombinere enkelte av Schaeffers lydobjektkonsepter med moderne kognisjonsforskning<sup>7</sup>, argumenterer Rolf Inge Godøy for at de mentale og kroppslige

---

<sup>5</sup> Dataspill er et viktig unntak, samt andre bruksmønstre hvor det spilles musikk i bakgrunnen

<sup>6</sup> Schaeffer selv foreslår segregering av lydobjekter gjennom det han kaller stress-ramheving (Schaeffer 1966:366), som relativt sett også kan sees på som betydelige endringer i lydintensitet (Godøy 2006:151). Dette skiller seg ikke alt for mye fra den ‘brudd på stillhet’-tolkningen som jeg har gått for.

<sup>7</sup> Godøy samler likhetstrekk mellom flere fagtradisjoner til å danne et fenomen han kaller motormimetisk kognisjon (Godøy 2006:155), og som danner grunnlaget for Jensenius’ handling-lyd-slektskap (Jensenius 2008:19)

bevegelseskonseptene i den lydskapende bevegelsen er noe vi alle kjenner oss igjen i, og at betydningen ligger på et så basalt plan at det ikke er nødvendig å foreta noen redusert lytting for å identifisere lydobjektets bevegelse:

“there is a gesture component embedded in Schaeffer’s conceptual apparatus which is on a more general and basic level than [the casual level] that the principle of reduced listening is supposed to lead us away from.” (Godøy 2006:154)

Det fruktbare med denne påstanden i forhold til analysen min, er at selv om jeg analyserer lyder ved hjelp av redusert lytting, vil den lydskapende gesten jeg finner også kunne gjelde for andre mer hverdagslige lytteperspektiver. Betydningen blir kanskje annerledes avhengig av lytteperspektiv, men ved å lytte etter lydskapende gester finner jeg i det minste en fellesnevner som gjør det lettere for meg å finne trekk i den historiske utviklingen til systemlydene.

### 2.2.3 Lydobjektets typologi og morfologi

Nå som jeg har etablert at den lydskapende bevegelsen til lydobjektet er en nyttig fellesnevner, passer det bra å se nærmere på hva slags kvaliteter ved lydobjektet som jeg kan lete etter i analysen. Schaeffer har laget noen diagrammer med slike kvaliteter som strekker seg over flere sider (Schaeffer 1966), men Godøy (2006) forenkler modellen og foreslår følgende inndeling: *typologi* og *morfologi*.

Typologiske kvaliteter finner man ved å studere lydens omhyllingskurve (Godøy 2006:153), og i lydteori er en omhyllingskurve en måte å se forandring i lydnivå over tid. Typologien til lydobjektet kan deretter ifølge Godøy ([loc.cit](#)) spores tilbake til tre kategorier av lydskapende bevegelser:

- Impulsiv – rask intensitetsøkning før lyden ebber ut – tilsvarer en slagbevegelse
- Utholdt – statisk lydnivå – tilsvarer en flytende bevegelse
- Iterativ – raskt gjentatte impulser – tilsvarer en friksjonsbevegelse

En impulsiv typologi kan forekomme i et piano hvor en hammer slår en streng, en utholdt typologi kan forekomme i et fiolin hvor lyden opprettholdes så lenge buen fortsetter å strykes

over strengene, mens en iterativ typologi kan forekomme når noen spiller på et vaskebrett. Grensen for hvor vi tolker en serie impulsive bevegelser som en iterativ typologi, går ved det punktet hvor det høres ut som om lydene glir inn i hverandre (Jensenius 2008:25). Fra denne tredelingen av lydobjektets typologi forstår jeg at ved å studere omhyllingskurven til en systemlyd, går det an å finne ut av typologiske kvaliteter ved lydobjektet. Derfra får jeg en pekepinn på hva slags lydskapende bevegelse lydobjektet har, som igjen kan si noe om hva slags handling som kan knyttes opp til lyden.

Morfologiske kvaliteter sier ikke så mye om den lydskapende bevegelsen, men er snarere en beskrivelse av mer interne trekk som klang og tekstur (Godøy 2006:153). Disse kvalitetene kan si noe om lydobjektets form, masse, kornethet og ganglag (ibid.:154). For eksempel kan tonehøyde beskrive massen til et lydobjekt ut fra en forestilling om at lyse toner er lette og mørke toner er tunge.<sup>8</sup> Et annet eksempel er hvis klangen endres, som da tilføyer en bevegelseskomponent i tillegg til lydobjektets typologi (loc. cit). I et slikt tilfelle vil man for eksempel kunne spore et nervøst ganglag ut fra en klangfarge som raskt svinger mellom to ulike tilstander. Imidlertid blir det vanskelig å finne konkrete betydninger i lydobjektet fra slike enkeltstående trekk. I situasjoner hvor jeg er på jakt etter betydning, kan det være fruktbart å bruke begreper fra semiotikken.

## 2.3 Multimodal semiotikk

For å finne systemlydenes betydning kommer jeg til å foreta en semiotisk analyse, og dette er i likhet med Schaeffers lydobjekter en metode som fokuserer på interne betydningsmønstre. Imidlertid vil jeg benytte meg av nyere teorier og begreper som tar for seg hvordan betydning strømmer mellom ulike representasjonsformer, og kan samles under begrepet multimodal semiotikk. Den metoden jeg vil bruke hører til medievitenskap, men terminologien kommer opprinnelig fra lingvistikk.

### 2.3.1 Det semiotiske tegnet

Ferdinand de Saussure delte ordene i språket vårt inn i to komponenter: En formkomponent (signifikanten) og en innholdskomponent (signifikatet) (Storey 2006:87). For ordet 'lyd' vil signifikanten være skriften 'lyd' du nettopp leste, mens signifikatet er ideen 'lyd' hvor vi

---

<sup>8</sup> Dette er lettere å slå fast hvis man trekker inn romlighetsmetaforer, som jeg går gjennom i Kapittel 2.4

forestiller oss noe hørbart. Sammen danner disse to komponentene et tegn, de *signifiserer* 'lyd'. Hvis vi ser for oss en systemlyd som et tegn, vil signifikanten kunne være lyden slik den høres, mens signifikatet vil kunne være en idé som lyden er ment å formidle. Forholdet mellom signifikanten og signifikatet er ifølge Saussure arbitrært (Storey 2006:87), altså er det kulturelt bestemt gjennom konvensjoner. Et brukergrensesnitt er fullt av slike tegn som er bestemt av konvensjoner, for eksempel er det ingen direkte sammenheng mellom signifikanten 'lukkerute' i et vindu og signifikatet 'avslutt'.

### 2.3.2 Tegnet i annen potens

'Et bilde sier mer enn tusen ord'. Det er vanskelig å motsi et så godt etablert uttrykk i språket vårt, og uttrykket regnes som sant fordi vi anser et bilde som et enkelt tegn som kan romme flere betydninger avhengig av hvem som ser på bildet. Mens vi lett forstår et bilde som en abstrahering av noe fra virkeligheten, er vi vant med at lyd alltid kommer fra en konkret lydkilde. Denne oppfatningen blir raskt problematisert når vi introduserer høyttaleren, en finurlig oppfinnelse som kan reprodusere lyd. Når den lyden som skal representeres ikke nødvendigvis har noe samsvar med en gjenstand eller hendelse i virkeligheten (som lett er tilfellet i en synthesizer eller datamaskin), kan lyder i likhet med bilder også regnes som en abstrahering. Da kan det være interessant å se på forskningen som er gjort på hvordan betydning kan strømme mellom bilder og skrift, og om et slikt forhold også kan gjelde for lyd i dens samspill med andre medietekster.

### Myten og konnotasjon

Roland Barthes ønsket å komme med "en ideologikritikk som gjelder språket i den såkalte massekultur; [...] et første forsøk på å demontere dette språket semiologisk" (Barthes [1957] 2002:3). Brukergrensesnitt er nødt til å tilpasses massene for at mange mennesker skal forstå dem, så jeg regner brukergrensesnitt for å være en del av en massekultur, slik Barthes bruker begrepet. Han introduserer teorien om *myter*, som ikke er gjenfortalte historier slik vi tradisjonelt forstår termen, men heller noe i vår kulturelle bagasje som "formidler falske selvfølgeligheter" (Lending 2002:XII). Slike falske selvfølgeligheter kan være lurt for meg å studere, slik at jeg ikke havner i fella med å anta at betydningen som jeg finner i en lyd, umiddelbart forstås av alle leserne.



Når noe 'gir lyd fra seg' kan vi tolke det tegnet som at noe produserer lyd. Imidlertid vil en leser raskt kunne trekke en annen naturlig tolkning av dette tegnet, nemlig at noe eller noen endelig tar bladet fra munnen. En slik tolkning får vi ifølge Barthes ([1957] 2002:295) ved å gjøre om tegnet 'lyd' til en signifikant og kombinerer den med et signifikat fra vår kulturelle bagasje, som i dette tilfellet er 'slutt på stillhet'. Spranget hvor et tegn igjen kan danne et nytt tegn er noe Barthes kaller *konnotasjon* (Barthes 1977b:37). "Konnotasjon setter således Saussures semiologiske system i annen potens" (Lending 2002:XXIII). Innen medievitenskap er det lett å se hvordan myten, hvor tegn hentes fra en felles kulturell bagasje, vil kunne si noe om en felles tankeretning eller ideologi:

"Et budskap vil *alltid* vise til, eller konnotere, sekundære, skjulte eller uklare betydninger, og nota bene: Det er dette ikke-umiddelbart tilgjengelige nivået som er det mest reelle, mest virksomme og det farligste – som reservoar for den ikke-erkjente ideologi som opprettholder samfunnets status quo." (Lending 2002:XXII, min utheving)

Fordi den kulturelle bagasjen som brukes til å hente fram en myte varierer fra individ til individ, er det ikke sikkert at en myte vil kunne forstås av en utenforstående. Dette er viktig å huske på når jeg finner myter i analysen av systemlydene, hvor en myte som fungerer for meg kanskje ikke fungerer som argument for en annen leser. Imidlertid står denne leseren selv fritt i å finne sine egne konnotasjoner i systemlyden, men da vil funnene også variere fra mine.

### 2.3.3 Forankring og avløsning

Som nevnt ovenfor kan et tegn leses på flere måter, men tegn står sjeldent alene, og vi tolker dem likevel stort sett ganske likt. Tydeligvis kan det ligge noe i denne kombinasjonen av forskjellige tegn, og Roland Barthes har konsentrert seg om hvordan mening produseres i kombinasjonen av bilde og skrift. Han hevdet at mens det i eldre kulturer var en reduksjon av mening fra skrift til bilde, har vi i nyere kulturer en *forsterkning* av mening fra den ene teksten til den andre (Barthes 1977a:26). Denne meningsforsterkningen kan også hevdes at går begge veier, ikke bare fra skrift til bilde, men også tilbake:

"I samspillet mellem flere forskellige tegnkæder skabes der en kompleks struktur, hvor betydning ikke blot produceres indenfor de enkelte kæder, men

også mellom kæderne – i kraft af redundanser og forskjelle mellem de enkelte tegn.” (Larsen 1988:27)

Fra sitatet ovenfor ser vi et eksempel på hvordan forskere etter Barthes har pekt på hvor innviklet betydningsstrømmer kan være, når det også oppstår betydning mellom tegnkjeder. I sine studier av bilde- og skriftkombinasjoner har Roland Barthes kommet med to aktuelle termer som beskriver hvordan enkelte betydninger velges framfor andre: *forankring* og *avløsning*.

## Forankring

“[Through the process of anchorage,] the caption [...] helps me to choose *the correct level of perception*, permits me to focus not only my gaze but also my understanding” (Barthes 1977b:39)

Forankring er når det er samsvar mellom innholdet i skrift og bilde, for eksempel at underteksten til et pressebilde nærmest gjentar hva som er avbildet, for å forsterke en spesiell betydning og skjule andre betydninger. Jeg tolker dette som at samsvaret i betydning ikke trenger å være innlysende i både bilde og skrift, men at det holder å kunne vise til at en betydning er synlig hos den ene, og at den understøttes av den andre. I den multimodale analysen min vil vi se et par eksempler på dette.

## Avløsning

“Text [...] and image stand in a complimentary relationship [...], and the unity of the message is realized at a higher level, that of the story, the anecdote, the diegesis” (Barthes 1977b:41)

Avløsning er når skriften tilfører nye signifikanter til bildet, om ikke nødvendigvis konnoterende og i annen potens som tilfellet er ved myten, så i alle fall i et komplimentært (nyskapende) forhold. Reklameplakater kan benytte seg av avløsning, særlig i tilfeller hvor det spilles på humor. Et bilde som viser en person som løper etter bussen med skriften “du luktet godt da du kom fram” bruker avløsning til å lirke fram en betydning om at denne deodoranten er effektiv for de som alltid er for seint ute. Dette er informasjon som ikke ble presentert i selve reklamebildet, men som ble introdusert gjennom tekstens avløsning i bildet.

Hvis man ser forankring og avløsning opp mot hverandre, har det vært argumentert for at forankring gjelder likhet og redundans i betydningsmønstre, mens avløsning gjelder interferens i betydningsmønstre (Larsen 1988:26). Dette betyr ikke at forankring og avløsning står i et ekskluderende forhold, i mange situasjoner kan avløsning forutsette forankring. En undertekst som sier 'Fotballsupporterne ga lyd fra seg' til et bilde av mennesker i fotballdrakt på en tribune, vil gjennom *forankring* få leseren til å tro at det ikke er nødvendig å tro at de var på et kostymeball. Samtidig vil skriften gjennom *avløsning* tilføre et element av at fotballsupporterne har brutt stillheten, og kombinasjonen av skrift og bilde danner da kanskje et tegn på en avgjørende dommeravgjørelse i fotballkampen.

### Kritikk mot å bruke forankring og avløsning

Roland Barthes' teorier begynner å eldes i forhold til moderne semiotikk, og de ble opprinnelig bare brukt til å beskrive sammenhengen mellom skrift og bilde (Barthes 1977b). Likevel finnes det forskere (Larsen 1988, Braaten og Toft ifølge Maasø 1995) som benytter seg av denne teorien til å analysere hvordan lyd forholder seg til levende bilder. Peter Larsen (1988) innrømmer at han bruker forankring og avløsning i en annen sammenheng enn det Barthes så for seg, men begrunner det på følgende vis:

“Når flere signifiant-kæder koordineres, etableres der et ladet betydningsfelt, – et felt hvor energier hele tiden *forskydes*, fordi identiske eller matchende signifikanter tiltrækker hinanden. I denne proces skabes der *fortætninger*, betydningsgrister, på tværs av kæderne, samtidig med at der også *artikuleres forskelle*. Set på denne baggrund bliver Barthes' to nøglebegreber, forankring og relæ, primært analytiske termer [...] forankring fokuserer på identitet, relæ på forskel.” (Larsen 1988:27)

Larsen ser altså forankring og avløsning som gode verktøy til å plukke fra hverandre kompliserte tegnkjeder på semiotisk vis. Arnt Maasø (1995) kritiserer forankring og avløsning for at de bare egner seg til å beskrive et styringsforhold, og at de forutsetter den oppfatning at en modalitet er underlagt en annen. Han etterlyser en bredere forståelse av lydens rolle opp mot levende bilder, et ønske om å se lydens tilleggsbetydning (Maasø 1995:24). Jeg anerkjenner denne kritikken, men jeg støtter meg til Larsens syn på forankring og avløsning som *verktøy* framfor å se dem som en restriktiv beskrivelse av lydens betydningsstrøm. En annen grunn til at jeg benytter meg av Roland Barthes' teorier, er fordi de stammer fra samme

periode som Schaeffers reduserte lytting, og at to strukturalistiske teorier kan fungere som en motvekt til de ferskeste teoriene som jeg benytter meg av i oppgaven. Den ene ferske teorien er handling-lyd-slektskapet som jeg tok for meg i begynnelsen av dette kapittelet, mens en annen er multimodal teori, som jeg nå skal ta for meg.

### 2.3.4 Multimodalitet

Multimodal forskning deler en tradisjonell oppfatning om at en *modalitet* er et meningsbærende medium (Van Leeuwen 1999:156), og at betydning eksisterer samtidig i flere modaliteter (Kress og Van Leeuwen 2001). Det er viktig ikke å forveksle modalitet med medium – skrift er fortsatt skrift både på papir og på en dataskjerm. Samme modalitet, men forskjellig medium. Det er også viktig ikke å forveksle modalitet med et semiotisk tegn; farge er en modalitet fordi den er en meningsbærer, mens ‘gul’ er et konkret tegn. Oppfatningen om hva som utgjør en modalitet er rimelig bredt, det kan til og med hevdes at plastikk som materiale er meningsbærende nok til å kunne kalles en modalitet (Kress og Van Leeuwen 2001:79). Med en så vid definisjon av modalitet, mener Theo Van Leeuwen at det trengs en snevrere definisjon av modalitet for å kunne analysere semiotikken til lyd: Graden av sannhet i en lydstrøm (Van Leeuwen 1999:180). Her menes ikke ‘sannhet’ som i lydens tilknytning til naturen, og heller ikke en objektiv sannhet, men mer som en troverdighet for lytteren gjennom en grad av semiotisk meningsforsterkning. Van Leeuwen (1999) deler denne sannhetsgraden inn i fire kategorier:

- Naturalistisk sannhet – lydens formmessige tilknytning til naturen, altså hvor realistisk og naturlig troverdig noe høres ut
- Abstrakt sannhet – lydens sublimitet og fortolkningskapasitet
- Skjematisk sannhet – lydens tekniske symbolverdi (mye brukt i måleutstyr), hvor man kan utlede mer eller mindre presise verdier ut fra en egenskap i lyden
- Sensorisk sannhet – lydens følelsesmessige innflytelse på lytteren (f.eks lyd i skrekkfilm)

Alle disse sannhetene eksisterer samtidig, som fadere på et miksebord (Van Leeuwen 1999:171), og det er vår kulturelle erfaring som gir utslaget for hvilken kategori som er tydeligst der og da. Når en eller flere av disse sannhetsgradene er til stede, opplever vi en *sannhetsforsterkning* (ibid.:172) som hjelper oss med å oppfatte en bestemt betydning fra lyden. Jeg tolker en slik sannhetsforsterkning som en mulig tilleggsbetydning som lyden fører med

seg, slik som det ble etterlyst i forrige del av dette kapittelet hvor jeg diskuterte kritikken mot forankring og avløsning.

Det er naturlig å tro at jo flere av de fire kategoriene som en lyd innehar, desto større sannsynlighet er det for at betydningen når fram til lytteren. Selv om jeg ser en mulig konflikt mellom det abstrakte og det naturalistiske ovenfor, vet jeg at lyd er et så komplekst fenomen at det ikke nødvendigvis er slik at det sublime ekskluderer det realistiske. Som vi kommer til å se, vil fordelingen av sannhetsforsterkning bli avgjørende når jeg skal prøve å finne betydningsstrømmen i min multimodale analyse av systemlydene. Det er en stund til vi kommer så langt, så jeg kommer til å friske opp disse termene med noen eksempler rett før den multimodale analysen.

### **Provenance og eksperimentelt meningspotensiale**

*Provenance* er et multimodalt begrep som dreier seg om de intertekstuelle forbindelsene knyttet til hvor vi mener et tegn stammer fra.<sup>9</sup> Dette begrepet ligner det Barthes omtaler som konnotasjon, men det nye er at konnotasjonen her skaper myter som eksisterer samtidig i flere modaliteter, og aldri monomodalt (Kress og Van Leeuwen 2001:75). For eksempel kan vi få lyst på iskrem av å høre munter musikk spilt av et elektronisk klokkespill, på grunn av den multimodale forbindelsen mellom musikkens karakter, barndomsminnet av å løpe etter iskrem bilen, og smaken av iskrem. På samme måte som med myten, er betydningen her avhengig av en kulturell bagasje som ikke bare representerer en ideologi i samfunnet, men som er langt mer individuell på grunn av de komplekse multimodale forbindelsene. Analyser som påviser provenance vil derfor benytte seg av individuell argumentasjon som ikke nødvendigvis leseren ville tenkt på, men hvor argumentasjonen kan være en analogi til leserens egne multimodale oppfatning. Noe av dette vil vi oppleve i min deskriptive analyse, og det samme vil også kunne inntreffe i den multimodale analysen.

*Eksperimentelt meningspotensiale* dreier seg om materielle egenskaper ved et tegn ut fra en oppfatning om hvordan vi kan gjenskape tegnet kroppslig (Kress og Van Leeuwen 2001:75). For eksempel kan vi ved å bruke ordet 'nydelig' i stedet for 'vakker' gi en assosiasjon om noe

---

<sup>9</sup> Provenance er engelsk for herkomst, slektskap eller avstamning. Kildene jeg har funnet som bruker denne fagtermen er alle på engelsk, og fordi en oversettelse av begrepet kan få en litt annen nyanse enn det opprinnelige, benytter meg derfor av det engelske begrepet.

mykt fordi D og G er mykere konsonanter enn K og kjennes myke ut å uttale i munnen.<sup>10</sup> Et annet eksempel kan vi finne i den kjente dusjscenen fra [Alfred Hitchcocks](#) skrekkfilm *Psycho* ([Hitchcock 1960](#)), hvor gjentatte lyse og intense fiolinslag høres og oversettes multimodalt til knivstikkene på lerrettet, med en smerte som vi også føler i kroppen hvis vi skulle prøve å gjenskape lyden ved å skrike. Eller lyden av negler skrapende langs en tavle, men da kommer vi mer inn på provenance igjen. Observante lesere har kanskje også tenkt at disse fiolinlydene kan være lydobjekter med en impulsiv typologi, som tyder på en slag- eller stikkbevegelse. Den største inspirasjonen bak denne teorien kommer imidlertid fra Lakoff og Johnsons metafor-teori ([Kress og Van Leeuwen 2001:77](#)), og i neste del av dette kapitlet skal vi se hvorfor.

## 2.4 Metafor-teori

Denne oppgaven heter ‘semiotiske systemlyder’, og jeg har så langt vært gjennom det lydmessige, deretter det semiotiske, og nå gjenstår det bare å sette det hele i *system* for at hele oppgavetittelen dekkles. Hvordan systematiserer vi? Jo, ved hjelp av metaforer. I boka *Hverdagslivets metaforer. Fornuft, følelser og menneskehjernen* ([Lakoff og Johnson \[1980\] 2003](#)) hevdes det at metaforene i språket vårt er helt sentrale i måten vi systematiserer verden på. “Vårt begrepssystem, som vi både tenker og handler ut fra, er grunnleggende metaforisk” (*ibid.*:7). Metafor-teorien hører til lingvistikken, men har også blitt utledet gjennom kroppslig persepsjonsforskning ([Johnson 1987](#)), ikke ulikt den økologiske tradisjonen som jeg tok for meg tidlig i dette kapitlet.

En metafor kjennetegnes ved at man forstår og erfarer et konsept ut fra et annet, enten gjennom naturlig eller kulturell erfaring. Et slikt hjelpemiddel kan komme godt med, når vi skal forstå de virtuelle objektene og hendelsene som presenteres i brukergrensesnittet til en datamaskin. Imidlertid er ikke metaforer noe som eksisterer bare i poesien og i brukergrensesnitt, [Lakoff og Johnson \(\[1980\] 2003\)](#) deler metaforer inn i tre overordnede kategorier:

---

<sup>10</sup> Dette er mitt eget eksempel hvor jeg bruker samme stoppkonsonant-bytte som et engelsk eksempel i ([Kress og Van Leeuwen 2001:77](#))

- Strukturmetaforer
- Romlighetsmetaforer
- Ontologiske metaforer

Jeg kommer ikke til å ta for meg strukturmetaforer, fordi de baserer seg på strukturalistiske teorier som diskurs og makt, og dette er et felt som jeg ønsker å holde utenfor denne oppgaven. Romlighetsmetaforer, også omtalt som orienteringsmetaforer, er når begreper knyttes til en bestemt plassering eller en bestemt bevegelse. Eksempler på dette er fornemmelsen av at ‘glad er opp’ og ‘trist er ned’ og hvordan metaforer kan reduseres til slike grunnleggende romlige uttrykk ([ibid.](#):17). Hovedargumentet til Lakoff og Johnson, at metaforer er grunnleggende i vår verdensforståelse, møter i dette tilfellet bistand fra kognitiv psykologi. Der er det blitt påvist at hippocampus, hjernens styringssentral for langtidshukommelsen, organiserer erfaringer på samme måte som den organiserer navigasjonsinformasjon ([Snyder 2000](#)). Tilsvarende er også navigasjon en hyppig nevnt ferdighet når man skal betjene et brukergrensesnitt. I tillegg til å se romlighetsmetaforer som nyttige for navigasjon, er et annet viktig poeng at vi liker å se romlighetsmetaforer i relasjonelle forhold, som opp/ned, inn/ut, forside/bakside og nær/fjern. Disse forholdene kan utnyttes når man skal designe systemlyder, særlig hvis man tenker på den lydskapende gesten til lydobjektet og dens forbindelse til romlige bevegelsesinntrykk ([Godøy 2006](#)). Det teoretiske begrepet ‘romlighetsmetaforer’ blir utledet gjennom en lingvistisk argumentasjon ([Lakoff og Johnson \[1980\] 2003](#)), men det har en beslektet term i kognitiv forskning: *image schemata* ([Johnson 1987](#)). *Image schemata* beskriver også en romlig forståelse, men da som ulike bevegelsesinntrykk som skapes og forsterkes av tilbakevendende mønstre i vår daglige erfaring med verden ([ibid.](#):29). Fordi argumentasjonen her bygger på menneskets erfaring med verden, har den et likt utgangspunkt som den økologiske persepsjonsforskningen som danner grunnlaget for bevegelsesinntrykk i handling-lyd-slektskapet.

Ontologiske metaforer er ifølge [Lakoff og Johnson \(\[1980\] 2003\)](#) menneskets virkemiddel for å materialisere og kategorisere abstrakte begreper. Disse metaforene kan videre deles inn i gjenstandsmetaforer, stoffmetaforer og beholdermetaforer. “Hendelser og handlinger begrepsliggjør vi metaforisk som gjenstander, aktiviteter som stoffer, tilstander som beholdere.” ([ibid.](#):32). Inndelingen begrunnes ut fra eksempler fra språket vårt, da særlig bruken av preposisjoner for å vise metaforisk tilhørighet, samt verb som brukes i

personifikasjon (ibid.:34). Det er verdt å merke seg at de overnevnte “hendelser, handlinger, aktiviteter og tilstander” alle er begreper som et brukergrensesnitt prøver å formidle. Vi ser dette igjen i terminologien: Knapp (gjenstand), meny (gjenstand), surfing (stoff: vann), mappe (beholder), skrivebord (beholder:bord/skuff). I analysen kommer jeg til å snakke om ‘tekstur’ – det er viktig ikke å forveksle dette med strukturmetaforer, selv om ordet ligner. Tekstur brukes i dette tilfellet som en stoffegenskap, og hører derfor hjemme under stoffmetaforer, morfologiske kvaliteter og materialitet.

## En brobygger

Jeg mener det er viktig å kjenne til metafor-teori før en analyse av systemlyder, fordi den er en brobygger mellom de ulike fagtradisjonene som jeg benytter meg av. Som vist i dette kapitlet, finner vi spor av romlighetsmetaforen både i den bevegelsestolkningen av lydobjektets typologi og i eksperimentelt meningspotensiale. Ontologiske metaforer finner vi tilsvarende i lydobjektets morfologi. For å komplettere den teoretiske brobyggingen, finner vi begge metafortyper i brukergrensesnittet som lydobjektene befinner seg i. Fordi metafor-teori har en slik brobyggende rolle, kan betydning som jeg finner via metoder fra medievitenskapen settes opp mot betydning funnet via metoder fra lydteori, og omvendt. Dermed blir det betydelig enklere for meg å få resultater når jeg skal analysere den multimodale betydningsstrømmen i systemlydene.

## Oppsummering

I dette kapitlet har jeg gjort rede for oppgavens analytiske teorigrunnlag, som involverer lydteori med bakgrunn i økologisk persepsjonsforskning, men også tradisjonell semiotikk, moderne multimodal teori og med metafor-teori som en overordnet paraply. Jeg har sagt at jeg vil benytte meg av en redusert lytting til å finne betydning ut fra lydens interne egenskaper. *Handling-lyd-slektskapet* tilsier at jeg skal kunne finne en sammenheng mellom lyden og handlinger i brukergrensesnittet via å studere bevegelsesinntrykk i lydobjekter. Lydobjektets *typologi* vil kunne skape et godt utgangspunkt for å se slike bevegelsesinntrykk i lyden, og typologien kan være impulsiv, utholdt eller iterativ. Lydobjektets *morfologi* kan gi supplerende informasjon om kvaliteter ved objektet som vi knytter bevegelsesinntrykket til.

Bevegelsesinntrykket vil knyttes til betydning gjennom semiotikk. Jeg har vært gjennom hvordan det semiotiske tegnet er satt sammen, at tegnet er basert på et arbitrært forhold



mellom signifikanten og signifikatet, og at tegnet av natur kan ha mer enn en bestemt betydning. Tegnet vil også kunne danne nye tegn gjennom *konnotasjon*, en prosess som er avhengig av leserens kulturelle erfaring. Denne konnotasjonen kan foregå multimodalt, da via et fenomen som kalles *provenance*. Ikke bare kan betydning oppstå på tvers av lyd, skrift og bilde, men ifølge *eksperimentelt meningspotensiale* også gjennom interne kroppslige inntrykk. For å kunne se sammenhenger i denne jungelen av betydningsstrømmer, går det an å studere hvordan betydning *forankres* eller *avløses*, men også hvordan ulike former for *sannhetsforsterkning* kan tilføre ekstra betydning gjennom en modalitet alene. Forankring, avløsning og sannhetsforsterkning kommer jeg ikke til å benytte meg av før i Kapittel 4, men de andre begrepene nevnt ovenfor vil være analyseverktøyene mine i neste kapittel, hvor jeg vil forsøke å finne betydning ut fra lydens interne egenskaper.

## 3. Analyse

---

**M**ed teorien fra forrige kapittel friskt i minne, passer det komme i gang med å analysere systemlydene. Før vi kommer så langt, må jeg presisere hvilke tekniske og kategoriske valg jeg har utført. Etter at disse valgene er gjort rede for, vil jeg først analysere lydene deskriptivt for å finne deres interne semiotiske betydning, og så bruke funnene mine til å besvare denne oppgavens første problemstilling, nemlig hvordan systemlydenes semiotiske trekk har utviklet seg gjennom analyseobjektene.

### 3.1 Tekniske valg og kategorisering

#### 3.1.1 Analyseverktøy

I tillegg til å anvende redusert lytting via hodetelefoner og studiomonitorer i et lydstudio, studerer jeg systemlydene ved hjelp av noen grafiske verktøy. Gitt denne oppgavebesvarelsens skriftlige, visuelle og ikke-auditive modaliteter gitt av papirets begrensninger, vil det være lettere å vise argumentasjonen visuelt enn om jeg bare skulle forklart leseren det jeg selv mener å høre. Til disse visuelle framstillingene har jeg brukt lingvistikkanalyseprogrammet PRAAT 5.1 ([Boersma og Weenink 2009](#)) til å lage bølgeformer og spektrogrammer av hver enkelt systemlyd. Bølgeformenes hovedformål er å påvise lydobjektets typologi, mens spektrogrammenes hovedformål er å påvise lydobjektets morfologi. I de neste avsnittene vil jeg forklare hvordan dette lar seg gjøre.

#### Bølgeformer

Bølgeformene viser lyd beskrevet som amplitudesvingninger over tid, og det vertikale tverrsnittet (enklere forklart, tykkelsen) er korrelert til den perseptuelle opplevelsen av lydstyrke. Jeg har brukt bølgeformenes evne til å vise endringer i lydstyrke for å få informasjon om systemlydens omhyllingskurve. Denne omhyllingskurven består av tre faser: Den første er *anslaget* (den innledende volumøkningen), den andre er *kroppen* (noenlunde statisk lydstyrke) og den siste er *halen* (hvor lyden ‘slippes fri’ og ebber ut). Jeg vil i hovedsak bruke lydens omhyllingskurve til å finne lydobjektets typologi, da visse lydskapende handlinger produserer

kjente mønstre i lydens energiutvikling over tid (Godøy 2006). Av denne typologien håper jeg å komme fram til en beskrivelse av hva slags bevegelse som kan ha skapt systemlyden.

De systemlydene som er i stereoformat har en bølgeform for hver kanal. Kanalene er plassert under hverandre. Slik blir det mulig å påvise bevegelse i stereobildet via lydstyrkeforskjeller mellom de to kanalene. For å få tydeliggjort detaljer i bølgeformene, har jeg normalisert dem slik at bølgeformene til de ulike systemlydene er omtrent like høye, selv om dette kan villedde leseren til å tro at systemlydene har samme lydstyrke. For å redusere sjansene for en slik villedning har jeg også notert ned lydfilens kraftigste amplitude, målt i antall desibel fra digital maksverdi (dBFS). Desibeltallet tolket alene er ikke spesielt interessant, da det endelige lydstyrken vil avhenge av hvor mye forsterkning det er i avspillingsutstyret, men det fungerer for å kunne se hvilke lyder som er kraftigere enn andre. Forskning på forholdet mellom lydobjektets typologi og egenskaper ved den lydskapende bevegelsen har påvist at forskjellige lydstyrker kan påvirke analyseresultater (Grassi og Burro 2003:50). Derfor syns jeg det er viktig å ha med en viss indikator på systemlydenes relative lydstyrke, selv om det overordnede formålet med denne analysen er å komme fram til semiotisk betydning.

## Spektrogrammer

Spektrogrammene viser lydens energi i ulike frekvensområder over tid. Den vertikale aksene beskriver frekvensene, mens den horisontale aksene beskriver tid. Jo mørkere spektrogrammet er for en viss frekvens, desto mer energi er det i den frekvensen på det tidspunktet. Det er ikke et poeng for meg å tyde tallverdier av spektrogrammene, men snarere å bruke spektrogrammets evne til å få et visuelt inntrykk av lydens klang. Ifølge Microsoft har systemlydene som jeg analyserer hørbare toner i frekvensområdet mellom 600-2000 Hz (Microsoft 2008:705), mens høyere frekvenser kan være overtoner som sier mye om lydens klangfarge, tekstur og fluktuasjoner over tid.

Jeg har brukt standardinnstillingene for spektrogrammer i PRAAT 5.1, men frekvensområdet er forstørret så det dekker de viktigste komponentene<sup>11</sup> Da det ikke er like nødvendig å se ørsmå tidsdetaljer i systemlydene som det er i en analyse av for eksempel språklyder, forlenger

---

<sup>11</sup> Opprinnelig frekvensområde 0-5 kHz, økt til 0-10 kHz. Som barn kan vi høre opp mot 20 kHz, men dette tallet reduseres etter hvert som vi blir eldre.

jeg tidsvinduet fra 0,01 til 0,02 sekunder.<sup>12</sup> Dette gir høyere oppløsning i frekvens på bekostning av lavere oppløsning i tid. Jeg gjorde dette fordi det ikke er like nødvendig å se ørsmå tidsdetaljer i systemlydene som det er i en analyse av for eksempel språklyder. På den annen side er det fordelaktig for analysen min å kunne se små detaljer i hørbare toner og klang, og derfor begrenser jeg meg til mer generelle trekk over tid.

Avslutningsvis i denne tekniske redegjørelsen ønsker jeg å komme med en presisering av min metode: det at jeg bruker grafiske verktøy, betyr ikke at de har vært utgangspunktene mine for å finne den semiotiske betydningen. Mitt viktigste verktøy er ørene gjennom redusert lytting. De grafiske verktøyene bruker jeg til å påvise hørbare trekk i visuelt presenterbar form.

### 3.1.2 Valg av lyder og deres kategorisering

En systemlyd i Microsoft Windows kan spilles av ved å knyttes til en lydhendelse. Med ‘hendelse’ mener jeg i denne sammenheng et metodekall fra Windows-rammeverket som programmerere kan benytte seg av når de utvikler et Windows-dataprogram, på engelsk kalt “system event” ([Microsoft 2008:20](#)). Det er viktig ikke å forveksle hendelser med ‘handling’ og ‘lydskapende handling’, da disse omtaler forholdet mellom lyd og bevegelse (se [Kapittel 2.2](#)). For noen kan det være lettere å forstå terminologien ved å se for seg de engelske termene, hvor ‘event’ og ‘action’ er ord som ligger lenger fra hverandre enn ‘hendelse’ og ‘handling’.

Det er mange systemlyder som følger med Microsoft Windows. Tilsynelatende har Microsoft gjort det enkelt for meg å avgrense hvilke hendelser som systemlydene hører til. Ved å gå inn i lyd-kontrollpanelet dukker det opp en oversikt over hvilke hendelser som kan ha lyder knyttet til seg og en oversikt over hvilke lyder som kan knyttes til disse hendelsene. Utvalget av hendelser har ikke forandret seg nevneverdig gjennom de ulike systemversjonene, så hendelseskategoriene og deres standardlyder er stabile nok til at jeg kan følge denne inndelingen i analysen.

Vi er vant med å kunne velge egne ringelyder på mobiltelefonene våre. Tilsvarende er det også mulig å ha sine egne lydoppsett på datamaskinen. Med et så stort antall mulige kombinasjoner av ulike systemlyder i Microsoft Windows, er det derfor nødvendig med en viss

---

<sup>12</sup> 0,02 sekunder tilsvarer 441 samples ved 22.05kHz og 882 samples ved 44.1kHz, som er de to samplingsfrekvensene som brukes i analyseobjektene mine.

avgrensning før jeg går i gang med analysen. Nedenfor står de lydhendelsene jeg vil analysere, og min vurdering bak valget av akkurat disse hendelsene. I hovedsak er det fire kriterier som ligger til grunn for utvalget mitt:

- Lydene og deres hendelser skal være uforandret fra den inndelingen de får etter en ren systeminstallasjon. Slik vil funnene mine kunne gjelde en Windows-PC slik brukeren forventer å møte den. Hvis jeg forutsetter at mange av oss ikke finner det for godt å bruke tid og krefter på å stille inn egne lydoppsett på datamaskinen, regner jeg da med at funnene mine om historisk utvikling blir mer relevante enn om jeg hadde analysert et spesialtilpasset lydoppsett. Standardoppsettet av lydhendelser i Windows som jeg har valgt å følge heter 'Windows Default'.
- Lydene skal representere hendelser som kan sies å være del av et hverdagslig bruksmønster. Denne oppgavens formål er å påvise semiotiske trekk ved systemlyder og se mønstre dem imellom, ikke å gi en fullstendig oversikt over systemlydene i Windows. Poenget med bruksmønsterkriteriet er å få luket ut mindre relevante lyder, som for eksempel den systemlyden som spilles når vi får en innkommende faks. E-post er et eksempel på noe som vi nordmenn i 2009 bruker mye mer enn faks, og jeg tror flertallet av leserne er interessert i å lese om lyder som de kjenner fra hverdagen.
- Lydene skal ha en historisk utvikling gjennom de ulike utgavene av Windows. Med 'utvikling' mener jeg at lydene må ha noe felles som både kan tyde på liten og stor utvikling. En ryddigere definisjon er at forholdet mellom hendelser og systemlyder skal ha blitt ivaretatt gjennom systemversjonene, slik at jeg ikke må blande sammen eller dele hendelser. Ved å se på den historiske utviklingen, ønsker jeg å tydeliggjøre at lyddesign delvis er bundet av tommelfingerregler og historisk konvensjon, men at det samtidig eksisterer semiotiske friheter.

- Lydene skal ha en viss multimodal forbindelse med andre samtidige handlinger. Da tenker jeg på handlinger i datamaskinens virtuelle verden (både medierte og eksekverte) og på handlinger i brukerens fysiske verden. Det bør selvfølgelig eksistere multimodale forbindelser disse imellom, for at jeg skal kunne forankre funnene mine til moderne medievitenskapelig forskning som bruker multimodal tegnteori. Dette kriteriet er også viktig for at jeg skal få plukket ut noen velegnede lyder som jeg kan studere betydningsstrømmen til i **Kapittel 3.3**

### **3.1.3 Vurdering av lydhendelser**

Med utgangspunkt i kriteriene jeg nettopp nevnte har jeg kommet fram til seks lydhendelser som jeg vil ta for meg i analysen:

#### **1) Critical Stop**

Lyden som skal representere den mest dramatiske hendelsen, når noe er gått alvorlig galt med maskinvaren eller at programvare avsluttes som følge av datakræsje. Dessverre inntreffer slike hendelser relativt ofte i dataverdenen, så det kan være interessant å se hvordan lyden spiller inn i meningsformidlingen i disse frustrerende situasjonene. Denne hendelsen har dessuten tydelige likheter mellom systemlydene på tvers av de ulike versjonene av Windows, samtidig som det finnes kontraster som skaper en historisk utvikling.

#### **2) Default Beep**

Dette er den lyden som de fleste programmer benytter seg av, både de som Microsoft har skrevet selv og hos tredjeparts-programvare. Denne systemlyden inntreffer når utviklerne enten ikke har visst hvilke handlinger som vil ligge bak en lydhendelse, eller at de av ulike årsaker ikke ønsker å programmere inn mer spesifikke metodekall. Default Beep er lyden som har overtatt rollen for den sterile og skarpe pipelyden som var vanlig å bruke i eldre PCer. Jeg mener det er interessant å se på denne hendelsen ut fra at den egentlig ikke skal representere noe spesifikt, og hvorvidt dette er en utfordring når jeg nettopp ser etter en spesiell betydning i systemlydene.

#### **3) Battery Critical**

Denne hendelsen er eksklusivt tilknyttet en tilstand i maskinvaren, ikke til programvaren eller brukeren. Jeg velger denne hendelsen for bevisst å se om det eksisterer noen multimodal

forbindelse mellom maskinens virtuelle verden og brukerverdenen. Den har også en noe overraskende utvikling, som analysen min vil gå nærmere etter i sømmene.

#### **4) Notification**

Dette er en hendelse som gjerne inntreffer etter oppstart av systemet eller etter oppstart av programmer. Imidlertid er det ofte en forsinket melding i forhold til selve oppstartshandlingen, fordi det som regel er bakenforliggende prosesser som har funnet det for godt å gi melding om noe. I og med at meldingen er forsinket i forhold til handlingen, ledsages den i de nyere systemene av en subtil og lite forstyrrende snakkeboble med en tekstbeskjed. Denne hendelsen er valgt for å belyse den multimodale sammenhengen mellom lyden og den delvis synkrone tekstbeskjedens medieuttrykk.

#### **5) Device Connect/Disconnect**

Denne hendelsen har to sider som representerer samme sak, at datamaskinen finner eller ikke lenger finner en ekstern enhet. Slikt kan forekomme både etter en menneskelig handling (koble til og fra) og etter en teknisk handling (kabelsvikt eller utstyr som endrer på/av-tilstand). Windows 2000 har ikke denne hendelsen, og utgår derfor av analysen. Til tross for at jeg har en lyd mindre å analysere, mener jeg denne hendelsen representerer et såpass viktig multimodalt aspekt fordi den kan ses som en forlengning av den handlingen vi gjør i forkant av at lyden inntreffer. Derfor mener jeg at den fortjener å analyseres på bekostning av et redusert historisk perspektiv.

#### **6) New Mail**

Lyden for ny innkommende e-post er ikke nødvendigvis del av et vanlig bruksmønster, da den er programspesifikk. I 2009 er det mange som sjekker e-posten sin i nettbaserte klienter som Hotmail, Gmail og SquirrelMail, framfor å bruke en tradisjonell e-postklient, og slike klienter vil typisk ikke benytte seg av denne lydhendelsen. Likevel har e-postklienter som Microsoft Outlook egne fordeler som gjør at en del av oss fortsatt sverger til dem. For disse personene vil denne lydhendelsen være kjente toner. I likhet med Device Connect, har denne lydhendelsen en del multimodale forbindelser. En interessant forskjell eksisterer imidlertid. Der hvor Device Connect tok for seg fysiske enheter, snakker vi her om e-post som er en virtuell erstatning for tradisjonelle brev. Dette kan være en nyanse som det er interessant å studere nærmere.

### 3.1.4 Viktige unntak

Det er mange lydhendelser som jeg ikke har tatt for meg, og av disse er det særlig to kategorier som er såpass vanlige at jeg bør begrunne hvorfor de ikke blir analysert.

#### 1) Realistiske representasjonsformer

Hendelsen 'Explorer Start Navigation' inntreffer når man går inn på en nettside, eller når man har klikket på fram- og tilbakeknappene. Denne kunne vært interessant å studere fordi den hermer etter museklikk-lyden og fordi den har en representasjonsform som står i direkte samsvar med lyden fra den fysiske verden. Imidlertid er lyden den samme gjennom alle tre systemversjonene, og den er svært kort. Det blir derfor vanskelig å anvende de nærlyttingsmetodene jeg har tatt utgangspunkt i, og det blir lite å si om historiske trekk, annet enn at det ikke har vært noen utvikling.

'Empty Recycle Bin' kan høres når man tømmer papirkurven. Denne lydhendelsen har i likhet med hendelsen nevnt ovenfor en svært naturtro representasjon, i den forstand at den høres ut som papir som krølles sammen på samme måte som vi kanskje ville gjort hvis vi kastet noe i papirkurven. Denne hendelsen har noen små historiske variasjoner i lydens uttrykk, men det er generelt for lite til at jeg klarer å peke på typologiske og morfologiske forskjeller. Dessuten ønsker jeg i denne oppgaven å konsentrere meg om semiotikken til lyder som vi normalt har problemer med å knytte opp til en spesiell betydning. Hvis jeg da skulle begynne å drøfte lydens opplevde realisme i forhold til hvilken handling vi egentlig hører, vil det kunne dreie oppgaven litt over i en annen retning. Andre forskere har drøftet dette temaet tidligere ([Gaver 1993a](#), [Jensenius 2008](#)), og jeg kommer til å ta opp noen av deres synspunkter i det kontekstualiserende kapittelet etter analysen.

#### 2) Lydklipp med komplekst musikalsk innhold

Når vi slår datamaskinen på, hører vi en liten musikksnutt ved navn 'Start Windows', og ved pålogging hører vi 'Windows Logon'. Tilsvarende hører vi 'Windows Logoff' ved avlogging eller avslåing av datamaskinen. Problemet som jeg ser ved å analysere slike musikkbaserte systemlyder, er at disse lydene er langt mer sammensatte enn de korte systemlydene jeg ellers har valgt å fokusere på. Selv om det sikkert er gjennomførbart å analysere komplekse lydsekvenser med de samme metodene som for kortere lyder, står resultatene i fare for å bli i beste fall unyanserte og i verste fall irrelevante. Det vil derfor bli mer omfattende å anvende



teorier om lydobjekter på dem, og desto mer omfattende å tolke mikronivå-informasjon som typologi og morfologi om til semiotiske tegn. I tillegg står jeg i fare for å måtte dra inn enda mer musikkteori, og denne oppgaven balanserer allerede på en egg mellom ulike forskningstradisjoner. Det finnes imidlertid eksempler på semiotiske analyser av slike kjenningsmelodier, for eksempel en analyse av de fire tonene som til sammen utgjør varemerket til databrikkeprodusenten Intel ([Tagg 2008](#)). Jeg kommer dessuten til å trekke inn noen relevante perspektiver om styrking av varemerker og branding når jeg kontekstualiserer funnene mine i [Kapittel 5](#).

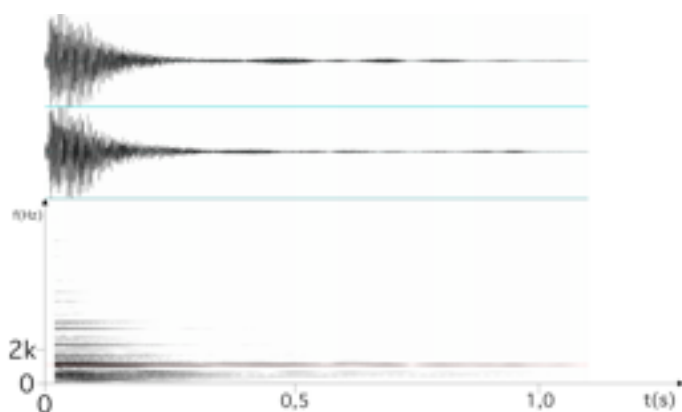
## 3.2 Deskriptiv semiotisk analyse

Nå som vi har gått gjennom inndelingen av lydhendelser til analysen, er det tid for å analysere lydene i detalj. I denne delen av oppgaven vil jeg lete etter følgende:

- Lydobjektets typologi. Om den lydskapende handlingen er impulsiv, utholdt eller iterativ.
- Lydobjektets morfologi. Trekk ved lydobjektet som sier noe om dets masse, størrelse, materiale, tekstur, form og plassering i rom.
- Hva slags bevegelsesinntrykk som kan knyttes opp til lydobjektene.
- Hva slags semiotiske tegn som systemlyden kan stå for.

Hver analyse av lydhendelser vil følge mønsteret gitt ovenfor, og for at det skal bli enklere å følge med og forstå hvordan hver enkelt lyd høres ut, anbefaler jeg at du som leser bruker lydmatrisen på nettsiden min ([Hjelmevold 2009](#)) aktivt mens du fordøyer innholdet i resten av dette kapittelet. Etter analysen av hendelsene kommer en oppsummering i [Tabell 3a](#), som også kan være nyttig å ha som referanse både underveis og i forkant av min besvarelse av problemstillingen i [Kapittel 3.3](#).

## Hendelse 1 av 6: Critical Stop



Figur 3a: Critical Stop (Windows 2000)

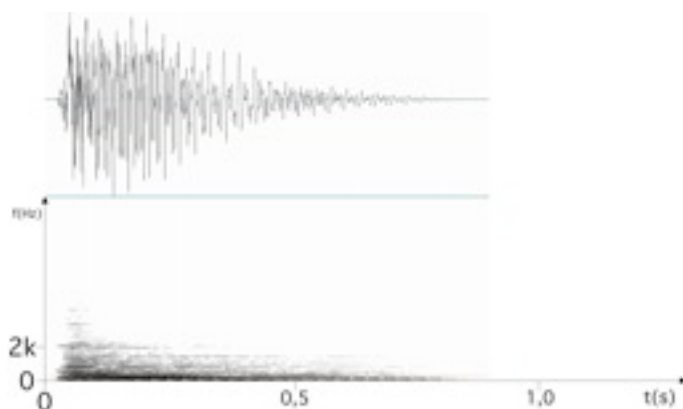
**Format:** 16-bit PCM 22,05 kHz Stereo

**Kraftigste amplitude:** -9,4 dBFS

**Varighet:** 1,1 s.

(Anslag: 0,01 s. Kropp: 0,1 s. Hale: 1 s.)

**Toneinnhold:** G-durakkord (G, H, D) spredt over mer enn en oktav



Figur 3b: Critical Stop (Windows XP)

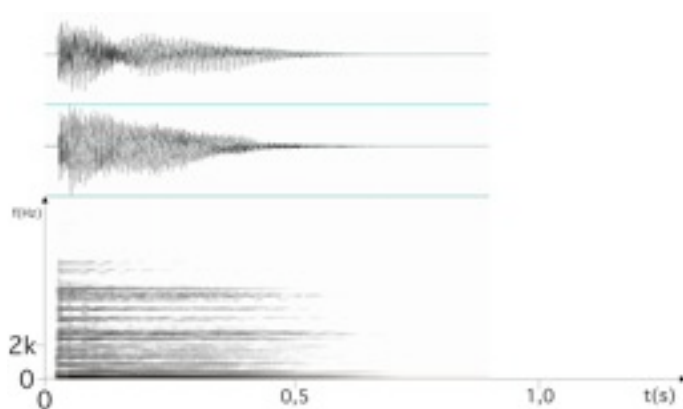
**Format:** 16-bit PCM 22,05 kHz Mono

**Kraftigste amplitude:** -3,2 dBFS

**Varighet:** 0,9 s.

(Anslag: 0,05 s. Kropp: 0,4 s. Hale: 0,4 s.)

**Toneinnhold:** En akkord med flere tonehøyder av C og en anelse G



Figur 3c: Critical Stop (Windows Vista)

**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -20,9 dBFS

**Varighet:** 0,9 s.

(Anslag: 0,05 s. Kropp: 0,2 s. Hale: 0,7 s.)

**Toneinnhold:** A-akkord spredt over flere oktaver

Vi begynner analysen av Critical Stop ved å vurdere lydobjektets typologi fra bølgeformene vist i [Figur 3a-3c](#). Bølgeformen til lyden i Windows Vista demonstrerer at det lett kan være uklart akkurat hvor de ulike fasene av omhyllingskurven begynner og slutter. Dette ser vi ved at halen, ifølge den ene kanalen, først begynner etter at lydstyrken har blitt svakere og så sterkere igjen, mens halen i den andre kanalen begynner allerede rett etter anslaget, uten noen kropp i det hele tatt. Det må utvises en del skjønn underveis når jeg avgjør omhyllingskurven, så det hele bør sees som en indikasjon framfor oppleste og vedtatte egenskaper. Når vi ser omhyllingskurven til lydene opp mot hverandre, er anslaget kortere og skarpere i Windows 2000, mens kombinasjonen av lengre anslag og lengre kropp gir Windows XP en mindre impulsiv og mer utholdt typologi enn de andre, om enn bare en svak forskjell. Hvis jeg skal anslå hva slags lydskapende handling som kunne skapt hvert lydobjekt, vil det for Windows 2000 og Vista være et hardt treff eller slag, mens det for XP er et mykere dunk, dytt eller brems. I det store og det hele har alle lydene en impulsiv typologi, men systemlyden til Windows XP er på grensen til å ha en utholdt typologi.

En mulig måte å avgjøre massen til et lydobjekt på, er ved å se om det tonale sentrum heller mot diskant eller mot bass (Schaeffer ifølge [Godøy 2006:153](#)). Massen til lyden i Windows 2000 er ifølge denne metoden langt mindre enn de to andre, på grunn av et tonalt sentrum lenger opp i diskanten. Windows XP har mye energi i bassen og er således veldig stor og tung, mens Vista har mindre energi i bassfrekvensene. Den låter fremdeles like mørk som XP fordi dens tonale sentrum også er i bassen, men i Vista fører mindre energi til mindre masse i lyden. Toneinnholdet beveger seg fra å være en treklang (G, H, D) i Windows 2000 til å bli en åpen kvint (C, G) i Windows XP til å bli en åpen oktav i Windows Vista (bare A-toner). For denne lydhendelsen har dermed toneinnholdet beveget seg fra smale treklanger til brede oktavakkorder. På den annen side følger ikke lydobjektets masse samme utvikling. Her har Windows 2000 minst masse, XP har mest masse, mens Vista ligger i midten, men den ligger nærmest massen til Windows XP.

### **Hvilket materiale?**

For å avgjøre materialiteten til de ulike lydobjektene, har jeg først og fremst brukt ørene, men det finnes også forskning som gjør at jeg kan påvise tilsvarende trekk ut fra spektrogrammet. Hvis vi begrenser oss til å de fire materialene tre, gummi, glass og metall, har vi ifølge [Giordano \(2003\)](#) følgende inndeling:

- Tre: Smal ansamling av kraftige frekvenser, spektralt sentrum heller mot bass
- Gummi: Smal ansamling av kraftige frekvenser, spektralt sentrum heller mot diskant
- Glass: Bredere ansamling av kraftige frekvenser, spektralt sentrum heller mot bass
- Metall: Bredere ansamling av kraftige frekvenser, spektralt sentrum heller mot diskant

Avvik kan forekomme på grunn av alle de komplekse forhold som til sammen danner en lyd, men denne inndelingen er grei å ha som et utgangspunkt.

Formen og teksturen til de forskjellige lydene bærer preg av en liten skarp, glatt og hard trebit i Windows 2000, en stor og litt knudrete klump i Windows XP og en litt flatere stor og myk klump, vegg eller plate av mer metallisk eller glassaktig karakter i Windows Vista. I Windows 2000 er frekvensspekteret relativt smalt, noe som er typisk for en del slagverksinstrumenter av tre. Skarpheten kommer fra det korte anslaget og den korte lydkroppen som avbrytes raskt. Det samme gjelder romklangen i lydens hale, som normalt ville fortsatt å dempes til det uhørbare, men hvor lyden rett og slett kuttes mens det fremdeles er mye romklangsenergi igjen. Lyden er glatt i dens fravær av kornethet i både kropp og hale. I Windows XP er lyden som nevnt litt mykere (men fremdeles et kraftig dunk), og det er en anelse støy i lyden. I spektrogrammet ser vi denne støyen som horisontale linjer med høy tetthet i frekvens, hvor linjene varierer i energi over tid. Støyen gir lyden en litt knudrete, moseaktig tekstur, og den tydelige teksturen skaper en følelse av nærhet i forhold til lytteren. Lyden kan ha form som en klump rett og slett fordi dens størrelse og ekstreme nærhet maskerer bort alt annet, og blir derfor noe mer udefinert, som en klump. I Windows Vista har lyden langt mer energi i overtonene, og får dermed et mer klirrende preg slik vi kjenner fra metall og glass. Størrelsen er kanskje like stor som lyden til Windows XP, men fordi den har mindre energi i bassen, låter den litt flatere.

## Nær/fjern

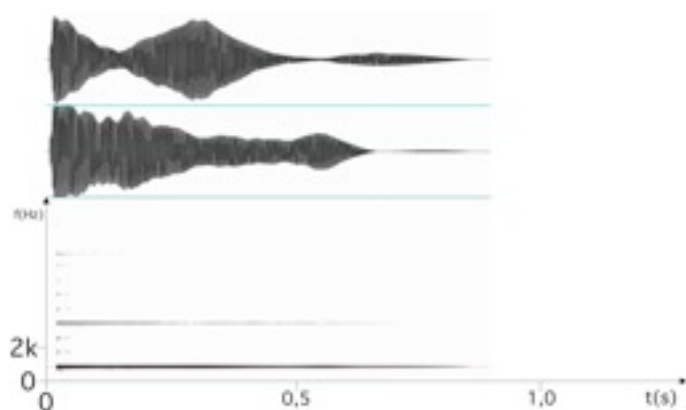
Romfølelsen er langt større i lyden til Windows 2000, takket være en lang hale med romklang. Kombinasjonen av lyd fra en liten og lett gjenstand i et stort rom gir et inntrykk av noe som begynner som et lite objekt, men når vi hører mer av det store rommet i etterklangen, blir romfølelsen forstørret. Slik åpnes lyden opp fra at vi hører et lite objekt til at vi hører et stort rom. Et annet virkemiddel som utnyttes i lyden til Windows 2000, er kontrasten mellom den nære lydkilden og den fjerne romklangen. Vi husker fra [Kapittel 2.4](#) hvordan beholdermetaforer er basert på et kontrastforhold av å være i noe eller utenfor noe ([Lakoff og](#)

Johnson [1980] 2003). Kontrasten mellom nær og fjern blir dermed en beskrivelse på en eller annen tilstand i datamaskinens virtuelle verden. Det er mindre romfølelse i lydene til de to senere operativsystemene, og langt mer følelsen av en stor og dominerende gjenstand som blokkerer alt annet. Særlig lyden til Windows XP låter påtrengende, da den store energien i bassen faktisk vil kunne ramme deg direkte i brystkassen hvis den spilles på høyttalere eller hodetelefoner som kan gjengi kraftige basslyder. Sannsynligheten for slike effekter øker når vi ser at lydstyrken til Windows XP bare ligger 3 desibel under maksnivå, som er vesentlig kraftigere enn de to andre lydene.

Felles for alle tre “Critical Stop”-lyder er at det brukes virkemidler som signifierer størrelse og dramatikk. Likevel brukes det forskjellige virkemidler for å formidle betydningen gjennom de tre ulike systemversjonene. Av disse er det kanskje skarpheten som er den egenskapen som i størst grad er felles for de tre lydene. Når vi kjenner skarphet på kroppen, dukker det opp tanker i bakhodet om farlige gjenstander, kanskje vekker det også overlevelsesinstinktet i oss. Disse multimodale assosiasjonene som skarphet gir, signifierer dermed dramatikk. Skarpheten er litt mindre framtrедende for lyden til Windows XP, men dette kompenseres for ved at denne lyden spilles av langt kraftigere enn de to andre. I Windows XP og Vista bidrar også bassinformasjonen til å øke lydens masse og romlige dominans, og oppfattelsen av en dominerende størrelse forsterker det semiotiske tegnet på dramatikk.

Den andre betydningen som er enkel å trekke ut av de tre lydene, er at de kan være skapt av et bevegelsesmønster, i dette tilfellet fra noe som stopper opp. Mye ligger i den impulsive kvaliteten til lydene, en typologisk kvalitet som de deler med slagverksinstrumenter. Imidlertid kan et slag eller et støt både skape bevegelse såvel som stoppe den. Denne egenskapen ved lydene er altså tydelig polysemisk, da den vel så godt kunne symbolisert det motsatte, nemlig noe som starter eksplosivt. I dette tilfellet kommer kanskje betydningen tydeligst fram i lyden til Windows XP, da den har et mykere preg som fra en bremseprosess, i tillegg til at den har en anelse friksjon innlemmet i lydens tekstur. Som om en bil skulle bremse, eller som om noe stort skulle bli satt ned på bakken.

## Hendelse 2 av 6: Default Beep



Figur 3d: Default Beep (Windows 2000)

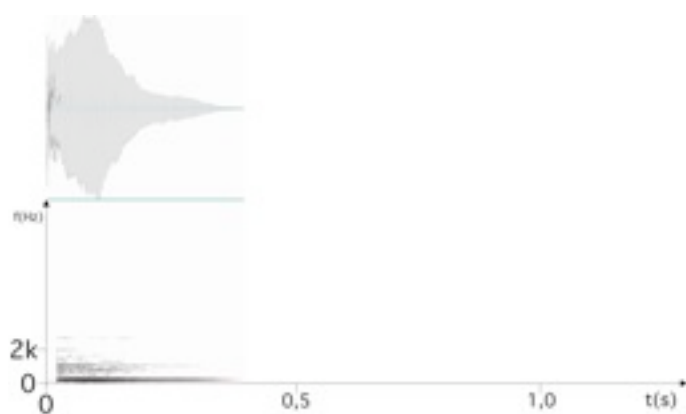
**Format:** 16-bit PCM 22,05 kHz Stereo

**Kraftigste amplitude:** -14,5 dBFS

**Varighet:** 0,9 s.

(Anslag: 0,01 s. Kropp: 0,6 s. Hale: 0,3 s.)

**Toneinnhold:** G



Figur 3e: Default Beep (Windows XP)

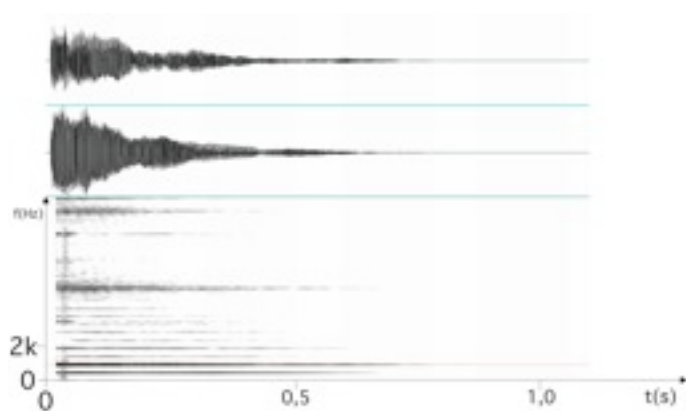
**Format:** 16-bit PCM 22,05 kHz Mono

**Kraftigste amplitude:** -11,7 dBFS

**Varighet:** 0,4 s.

(Anslag: 0,002 s. Kropp: 0,1 s. Hale: 0,3 s.)

**Toneinnhold:** H



Figur 3f: Default Beep (Windows Vista)

**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -22,7 dBFS

**Varighet:** 1,1 s.

(Anslag: 0,01 s. Kropp: 0,3 s. Hale: 0,8 s.)

**Toneinnhold:** A-akkord spredt over flere oktaver

Alle lydene har et svært kort anslag. Lyden fra Windows XP er kortest i alle ledd, Windows 2000 har en lengre kropp, mens Windows Vista har en lengre hale. Ingen av disse forlengelsene er vesentlige i forhold til lydens karakter, så typologien er derfor impulsiv for alle de tre lydene. Variasjonene i omhyllingskurven har derimot en litt større innflytelse på oppfatningen av massen til hvert lydobjekt. Til tross for at det tonale sentrum til lyden i Windows 2000 er nærmere diskanten enn i Windows XP, er den like tung eller kanskje tyngre fordi den bevarer energien sin lenger. Den ligger et stykke opp i diskanten, så den er ikke spesielt tung, men den er samtidig større enn i Windows XP. Den forlengede halen til lyden i Windows Vista inneholder forsterkede overtoner, og ved å opprettholde energien i halen, forsterkes de frekvensene som gir et svevende, stigende og nesten sakralt preg på lyden. Alle tre lyder er såpass impulsive at det virker som om de har store mengder potensiell energi som plutselig utløses, og det er vanskelig å spore noen forutliggende årsak til energiutløsningen, i hvert fall ut fra bevegelsesmetaforer.

Default Beep-lydene befinner seg i tre forskjellige tonale plasseringer: midten, lavt og høyt. Windows 2000 har lydobjektet plassert i midtsjiktet, hvor den mellomlyse tonen opprettholdes statisk, men svevende fordi den veksler i intensitet mellom høyre og venstre kanal. Kanskje som et vaiende flagg? Windows XP befinner seg derimot langt nede, som en sur bemerkning fra en student i salen. Det sure elementet kommer inn fordi det viser seg at gjennom alle analyseobjektene i Windows XP, bare finnes tonen H i denne lyden, mens de andre lydene i Windows XP har tar utgangspunkt i tonen C. I vestlig musikktradisjon er det lille sekundintervallet mellom disse to tonene noe som lett kan klinge surt i melodier og akkorder, men denne betydningen er nok kulturelt betinget, da intervallet kan ha en annen rolle i andre musikkulturer. Innunder samme vestlige musikkforståelse er H ledetonen i C-durskalaen, selv om dette forholdet er relativt. Tonen henger altså og venter i spenning på at noe skal skje: i musikken er dette å gå til grunntonen C, mens i brukergrensesnittet er det kanskje slik at maskinen henger spent i påvente av at brukeren gir en form for tilbakemelding? Så selv om tonen er plassert lavt i bassregisteret, har den musiskalsk sett likevel noen svevende elementer i seg. I Windows Vista befinner lydobjektet seg utvilsomt høyt oppe i forhold til de andre lydene, da dens tonale sentrum ligger langt oppe i diskanten. Det er ingen bassinformasjon, og dette forsterkes av de lyseste overtonene og fornemmelsen av bevegelse oppover. I tillegg er det også et romlig forhold i likheten i klangen mellom Vistas lyder for Critical Stop og Default Beep, det er som om samme instrument blir spilt, bare i et lysere register for Default Beep.

## Anonymitet

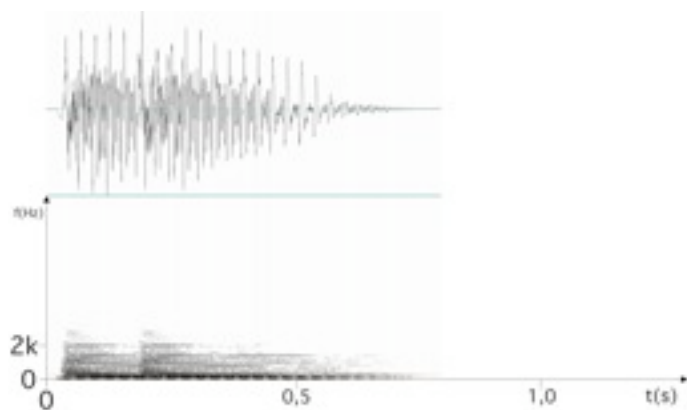
Teksturen til alle de tre lydene er glatt og ren, hvor dette er tydeligst i Vista-lyden på grunn av dens energi i mange diskantfrekvenser. Denne høyfrekvente energien tilfører en støy i lyden med så mange frekvenser at hvis vi her hadde snakket om lys, ville vi sett et hvitt lys. Denne lyse klangfargen kan hos enkelte lyttere i den vestlige verden skape en multimodal forbindelse mot religiøse høytider, som for eksempel julesanger. Som vi husker fra multimodal semiotikk, kalles en slik forbindelse for provenance ([Kress og Van Leeuwen 2001](#)). Glatt, ren og hvit er alle egenskaper som kan knyttes til sterilitet, noe som igjen kan knyttes til noe anonymt. Den sure og mørke tonen til Windows XP ødelegger nok noe i dette bildet, men den er i det minste fri for støy og annet grums i selve teksturen til lydobjektet. Alle tre lyder har noe metallisk ved seg, selv om XP heller mot treverk og Vista heller mot glass. Det metalliske kommer til uttrykk gjennom skarpheten i lydene, lyden av treverk kommer til uttrykk gjennom det litt hule og smalspektrede (se på bølgeformen til Default Beep i Windows XP), mens lyden av glass igjen kommer fra et lydbilde med høy intensitet i overtonene.

Lydene som hører til Default Beep i de tre systemversjonene bærer alle preg av at de skal være allsidige, og signifierer i all hovedsak en anonym oppmerksomhet. Et slikt uttrykk kan virke selvmotsigende, men lydene er ment å varsle brukeren om at noe har skjedd i datamaskinen, uten at de skal farge hendelsen med noen bestemt betydning. Et utropstegn vil være en tilsvarende signifikant i den visuelle semiotikken. Hvis vi følger den romlige metaforen om at opp er glad og ned er trist ([Lakoff og Johnson \[1980\] 2003](#)), ser det ut til at den romlige posisjoneringen av lydene i Windows 2000 og Vista vinkler denne betydningen litt mer til det positive enn det Windows XP gjør med sitt sure, mørke uttrykk. Uansett vinkling skinner det sterile uttrykket igjennom for alle lydene, så det er mulig at nettopp det sterile kan ha vært et trekk som Microsoft har strebet etter da systemlydene ble designet for denne hendelsen.



## Hendelse 3 av 6: Battery Critical

Windows 2000: Se [Figur 3d](#)



Figur 3g: Battery Critical (Windows XP)

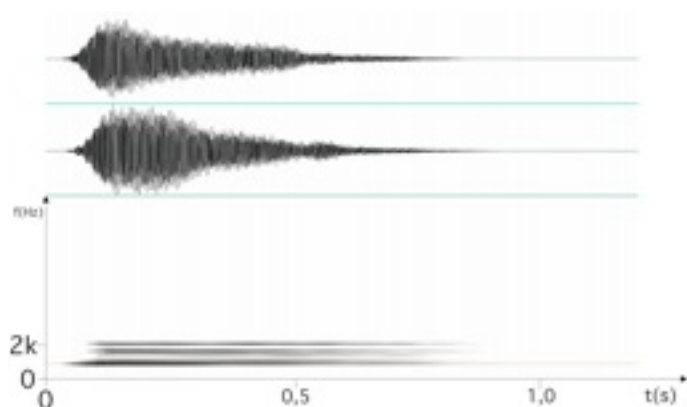
**Format:** 16-bit PCM 22,05 kHz Mono

**Kraftigste amplitude:** -0,2 dBFS

**Varighet:** 0,8 s.

(Anslag: 0,05 s. Kropp: 0,3 s. Hale: 0,5 s.)

**Toneinnhold:** C-akkord i flere oktaver spilt to ganger



Figur 3h: Battery Critical (Windows Vista)

**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -22,6 dBFS

**Varighet:** 1,2 s.

(Anslag: 0,1 s. Kropp: 0,1 s. Hale: 0,9 s.)

**Toneinnhold:** A, E, G (to kvintsprang) i glidende rask overlapping

For denne hendelsen er det en del gjenbruk av lyder; i Windows 2000 er det samme lyd her som for Default Beep, mens i Windows XP låter lyden som to anslag fra hendelsen Critical Stop rett etter hverandre. Lyden for lavt batteri i Vista er av en betraktelig annerledes karakter. Den er lys og har liten masse. Den lange anslagstiden gir den et mykt, streifende preg omtrent som et harmonisk vindkast. Lyden har en lang hale som ikke forandrer karakter i forhold til kroppen, så hvis disse segmentene regnes som en enhet, er kanskje dette det nærmeste vi kommer et lydobjekt med en vedvarende typologi innenfor de ellers så impulsive

analyseobjektene. Som vi ser av spektrogrammet inntreffer de lysere frekvensene senere og holdes lengre enn de mørke. Lyden beveger seg altså oppover i frekvensbildet, så den vedvarende bevegelsen er derfor også på vei oppover. En annen mulig bevegelsessammenligning er den hvor en finger strykes langs kanten på et krystallglass og så slippes. Bonusen ved en slik tolkning er de materielle og teksturelle egenskapene til glass som lydobjektet da får.

I og med at lyden som Windows 2000 bruker for å signalisere lavt batteri er akkurat den samme som den generelle feilmeldingslyden i Default Beep, følger den samme ubestemte betydningen med på lasset. Derfor blir det vanskelig å trekke fram noen betydning som er spesiell for denne hendelsen, og forholdet mellom signifikat og signifikant forblir derfor arbitrært og bestemt gjennom konvensjon. Hvis vi derimot studerer lydene fra de to andre systemversjonene, kommer vi lettere fram til en mer analog betydning.

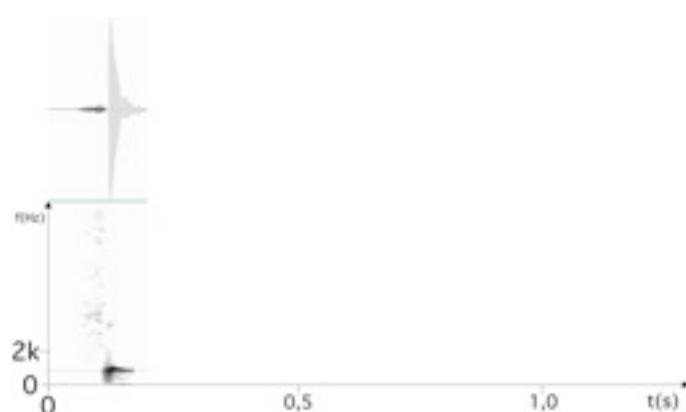
Selv om lyden for Windows XP inneholder de samme frekvensene som i lyden som brukes i Critical Stop og dermed har grunnlaget for en lignende betydning, introduseres det et rytmisk element som tilfører nye betydninger i lydobjektet. Jeg nevnte tidligere hvordan Critical Stop i XP bærer preg av nærhet, av og til i så stor grad at det rammer deg i kroppen. Vi husker fra multimodal teori at eksperimentelt meningspotensiale er når vi forstår et begrep gjennom hvordan det kjennes ut å skulle herme etter begrepet med kroppen ([Kress og Van Leeuwen 2001](#)). Hvis du får to hurtige slag i brystet, er det ikke mye som skiller disse rystelsene fra din eget hjertebank som du har blitt så vant med. Hjertet vårt er analogt med det å leve, så hvis vi gjennom eksperimentelt meningspotensiale forstår lyden som hjertebank, overfører vi dette til at lyden har noe med datamaskinens levedyktighet å gjøre. Bærbare datamaskiner fungerer ikke så godt uten strøm i batteriet sitt, altså signifierer lyden av hjertebank at maskinen ikke lenger er levedyktig inntil videre. At den arver betydningen av noe alvorlig og tungt. Batterier veier jo en del.

Det går an å finne en tilsvarende 'tapt levedyktighet'-betydning i lydobjektet for Windows Vista, men det er mulig at argumentasjonen henger i en tynnere tråd. Fordi lydobjektet har mindre masse, blir det vanskeligere for lyden å formidle alvor. Dette kunne vært gjort ved å ha noe dissonant skjærende i diskantområdet, men som vi ser av spektrogrammet, mangler det nesten all informasjon i overtonene. Ut fra de egenskapene jeg har funnet, klarer jeg å se en

betydning fra den romlige metaforen som dannes av at frekvensbildet beveger seg oppover. Enkelte vil kanskje tolke lyden som datamaskinens siste åndedrag, men da er det mer naturlig med en bevegelse nedover. En mer religiøs tolkning kan være at batteriet tar med seg datamaskinens sjel til himmelen, men dette anser jeg for å være en svært fri tolkning basert på en god dose humor. Dessverre har lydobjektet enten en glasstekstur eller ingen tekstur i det hele tatt, så da blir det ingen forbindelse med batteri-ideen gjennom lydobjektets tekstur, da ingen av dagens batterier er laget av glass.

## Hendelse 4 av 6: Notification

Windows 2000: se [Figur 3a](#)



Figur 3i: Notification (Windows XP)

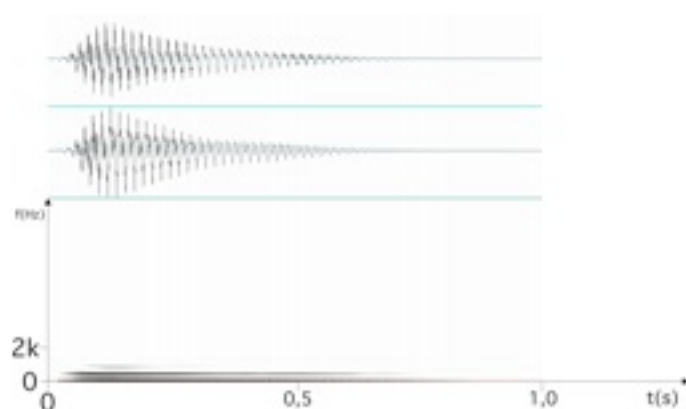
**Format:** 16-bit PCM 22,05 kHz Mono

**Kraftigste amplitude:** -18,2 dBFS

**Varighet:** 0,2 s.

(Anslag: 0,01 s. Kropp: 0 s. Hale: 0,05 s.)

**Toneinnhold:** G (høres bare i 0,05 sekunder)



Figur 3j: Notification (Windows Vista)

**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -12,9 dBFS

**Varighet:** 1 s.

(Anslag: 0,1 s. Kropp: 0 s. Hale: 0,9 s.)

**Toneinnhold:** En dyp og en høy A spilt samtidig med tre oktavers avstand

Igjen resirkuleres en systemlyd fra en annen lydhendelse, i Windows 2000 brukes lyden fra Critical Stop også til hendelsen Notification. Lydobjektet har derfor de samme kjennetegnene som for Critical Stop. Det skarpe kommer igjen, men denne gangen er det måten rommet åpner seg i lydobjektet som passer i forhold til det som lydobjektet skal bety. Når lydhendelsen Notification inntreffer, kan den medfølgende meldingen fortsette å ligge framme uten at du er forpliktet til å reagere med en eller annen handling. Så i stedet for at betydningen ligger i lydobjektets skarphet, ligger betydningen heller i romforståelsen som skapes av romklangen i halen til lydobjektet, som et tegn på åpenhet og tilbakemelding.

Typologien til Notification-lyden til Windows XP konfronterer meg med et lite problem. Jeg nevnte i [Kapittel 2.2](#) at jeg valgte å se bort fra segregering av lydobjekter, fordi korte systemlyder kunne anses som egne perseptuelt atskilte entiteter. Denne påstanden utfordres i systemlyden til XP, hvor det ikke er sikkert om den første halvdelen av lyden er et lydobjekt med utholdt typologi, eller om den bare er anslaget til det større lydobjektet. Det argumenteres for at lyden er så kort at komponentene uansett flyter over i hverandre. På samme måte som vi syns film er en jevn bevegelse når vi kommer opp i noen titalls bilder vist i sekundet, blir lyd også maskert i tid.

Det kan hevdes at frekvensinnholdet mellom de to delene er så forskjellig at de to halvdelene umulig kan henge sammen, men jeg mener at vi her er vitne til to lydobjekter som er del av samme handling. Hvis jeg forutsetter at andre halvdel av lyden er en impulsiv lyd som er skapt av en slagbevegelse, blir første halvdel tilhørende det som kalles “exciter” ([Grassi og Burro 2003:49](#)). Denne første halvdel av den lydskapende handlingen er når energi bygges opp og gir oss en forventning om at det vil komme en impulsiv lyd, som for eksempel når vi hører friksjonen som vinden skaper mot armene før vi slår hendene sammen for å fange en flue. Andre halvdel av lyden har et veldig kort anslag, ingen kropp og kort hale, og er impulsiv. Handlingen som helhet er derfor impulsiv, fordi den er en del av den samme lydskapende bevegelsen. Her kommer det imidlertid inn en viktig nyanse: I overgangen mellom spenningsoppbyggingen og -utløsningen er det i et lite øyeblikk ren stillhet. Dette stemmer ikke overens med måten energi overføres i en direkte slagbevegelse. Den bevegelsen som vi er vitne til her, er mer som en type sprekking. Se for deg luften som fyller en ballong opp til et punkt hvor den bare såvidt klarer å holde på energien, før ballongen plutselig sprekker.

Typologien til Windows Vista er langt fra like innviklet som det jeg beskrev om Windows XP ovenfor, her snakker vi om et lydobjekt uten kropp med et anslag som er lengre enn de to andre. Vi snakker om en impulsiv slagbevegelse her også, men det lange anslaget gir inntrykk av en langsommere bevegelse, kanskje mer en forbipassering eller streifing i stedet for et slag eller sprekking som vi har i de to andre lydene.

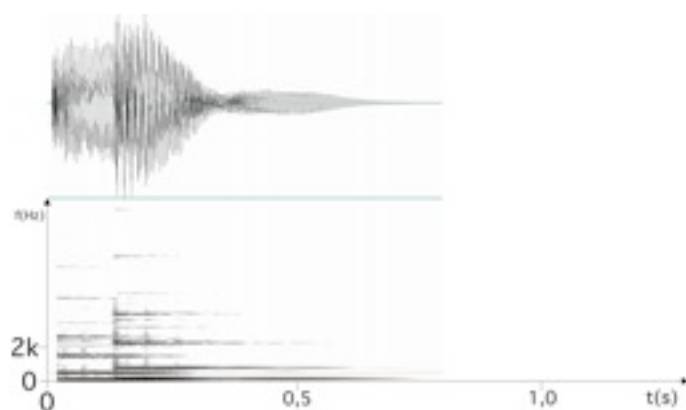
Som vi husker da jeg analyserte Windows 2000-lyden under 'Critical Stop'-hendelsen, kom jeg fram til at lydobjektet hadde en morfologi som tydet på en liten og hard gjenstand av tre, plassert i et stort rom. Morfologien til XP og Vista innunder den nåværende lydhendelsen er ganske annerledes. Vi hører ikke noen romklang som skulle tyde på at lydene befinner seg i noe rom, snarere befinner lydene seg helt inntil kroppen vår. Den hørbare tonen i systemlyden til XP er veldig konsentrert og ligger ikke helt nede i bassen, så den minner mer om gummi enn om treverk. Av samme grunn er massen rimelig liten. I systemlyden til Vista er det større spredning mellom de hørbare tonene, og det tonale sentrum er langt nede i bassen. Dermed blir dens materialitet noe stort og tungt, kanskje metallisk. Hverken XP eller Vista-lydene har en kropp og hale med noe særlig informasjon i diskanten, så det er vanskelig å få noe inntrykk av tekstur. Alternativet blir derfor at lydobjektet får en glatt overflate. For å oppsummere morfologien til lydobjektene, kan Windows XP være en liten og lett gummiaktig glatt gjenstand plassert inntil kroppen, mens Windows Vista kan være en stor og tung glatt metallisk gjenstand som også befinner nær kroppen.

Med et så stort sprik innenfor morfologiske trekk, virker det som om betydningen i disse lydene står i fare for å bli like kaotisk. For å komme fram til hva betydningen kan være, vil jeg studere de inntrykkene som skapes av lydobjektene romlige posisjon og bevegelse. I Windows 2000 begynner vi med et lite objekt som åpnes opp, og som nevnt ovenfor kan lyden på denne måten bli et tegn på at datamaskinen venter på tilbakemelding. Det er mulig det går an å se en lignende betydning i lyden til Windows XP fordi bevegelsen tyder på at et objekt åpnes i den forstand at det sprekker. Jeg stiller meg tvilende til en slik tolkning, fordi gjenstander som sprekker, sjelden etterlater særlig annet enn en samling med triste rester av det som var. En betydning jeg synes virker mer naturlig ut fra bevegelsesinntrykket, er overraskelse. Først hadde vi en bevegelse av energi som bygget seg opp, så kom det overaskende utbruddet hvor energien ble utløst. Lydobjektet er plassert nær lytteren, og siden materialet låter som gummi, er inntrykket av en tyggegummiboble som sprekker ganske nærliggende.

Når det gjelder betydningen av systemlyden til Windows Vista, er det vanskelig å komme fram til noe konkret. Det nærmeste jeg kommer er å tolke lyden som en et tegn på en bortgjemt bemerkning. Utgangspunktet for denne tolkningen kom da jeg sammenlignet denne lyden med Vista-lydene for Critical Stop og Default Beep. Alle lydene lå veldig likt, men denne var spilt lenger ned i bassen enn de to andre, den hadde lengre anslag, og den hadde lite eller ingen informasjon i diskanten. Klangmessig høres det mer eller mindre ut som om denne lyden er en utgave av de to andre lydene, men hvor all skarphet er filtrert bort. Det høres nesten ut som om lyden er blitt forsøkt gjemt bort, men at den 'tar en tur innom lytteren' på grunn av det lange anslaget. Lydobjektets morfologi som tilsier at vi hører en stor og tung glatt metallgjenstand på nært hold, klarer jeg ikke å koble opp mot noe i den betydningen jeg har funnet. Eventuelt går det an å se denne morfologien som en kontrast opp mot morfologien til Critical Stop i Windows XP. Der var det også en stor og tung gjenstand, men den var myk og knudrete i stedet for hard og glatt. Betydningen der var obstruksjon, så hvis jeg snur resonnementet på hodet, kan vi kanskje forvente oss at betydningen her skal være noe som ikke er i veien for det vi vil holde på med. For at lydobjektet ikke skal være i veien, gjemmes det ved å grave det ned i bassen og filtrere bort alle tegn på skarphet. Jeg føler riktignok at jeg må presisere, at i dette resonnementet har jeg vært noe uforsiktig ved å blande egenskapene glatt/myk som egentlig beskriver to forskjellige fenomener (friksjon/elasticitet). Derfor ser jeg betydelige mangler i å tolke betydningen til Notification i Windows Vista som et tegn på en 'bortgjemt bemerkning', men det er det mest presise jeg kan beskrive ut fra lydens interne egenskaper.

## Hendelse 5 av 6: Device Connect

*Windows 2000: Ingen slik hendelse.*



**Format:** 16-bit PCM 22,05 kHz Mono

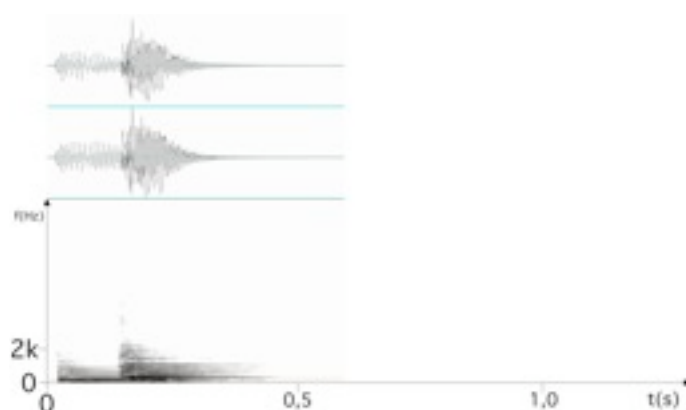
**Kraftigste amplitude:** -14,9 dBFS

**Varighet:** 0,8 s.

(Anslag: 0,02 s. Kropp: 0,2 s. Hale: 0,6 s.)

**Toneinnhold:** C som hopper en kvint opp til G

Figur 3k: Device Connect (Windows XP)



**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -14,2 dBFS

**Varighet:** 0,6 s.

(Anslag: 0,03 s. Kropp: 0,2 s. Hale: 0,3 s.)

**Toneinnhold:** A som hopper en kvint opp til E

Figur 3l: Device Connect (Windows Vista)

Lydene i denne hendelsen skiller seg ut fra de forrige hendelsene, fordi de tilfører et snev av melodi i stedet for bare å være et kort og statisk anslag. Vi skal snart se hvilke fordeler dette medfører for meningsformidlingen, men vi begynner som vanlig med å studere typologien til lydobjektene. Her har vi korte anslag i lydene til begge systemversjoner. Anslaget gjentas deretter med en annen tone enn den første. Denne slags typologi er på grensen til det iterative, men fordi lydene består av to forskjellige toner på forskjellig tidspunkt, mener jeg at de ikke flyter nok inn i hverandre til å kunne være ett iterativt lydobjekt. I stedet blir hver av systemlydene bestående av to impulsive lydobjekter med kort anslag, kort kropp og lengre

halen. Halen er lengst i systemlyden til Windows XP, og vi skal snart se hvordan dette forlenger den bevegelsen som ligger i samspillet mellom de to anslagene. Før vi kommer så langt, vil jeg nevne de morfologiske trekkene ved systemlydenes lydobjekter.

Lydobjektene i begge systemlyder har ganske stor masse, da det tonale sentrum befinner seg noenlunde i samme frekvensområde i bassen. Når det gjelder lydens materialitet, høres det ut som XP består av metall, fordi den har såpass stor spredning i frekvens. Vista-lyden er langt mattere, som om den skulle være mer gummiaktig. Det gummiaktige klarer jeg imidlertid ikke å påvise ved hjelp av frekvensspredning i spektrogrammet, men i stedet ved hjelp av den korte halen. Begge lyder har et spor av romklang, men det høres ut som de er plassert relativt nær lytteren.

Vi har i disse systemlydene i alt fire ulike hørbare toner, en C som går til en G og en A som går til en E. Selv om tonene ikke er de samme, er toneintervallet en kvint i begge systemlyder, og det er den dypeste tonen som spilles først. Jeg vil her hevde at det sentrale i betydningen til disse lydene er det forholdet som oppstår i kontrasten mellom to ulike toner. For å kunne argumentere for dette, henter jeg fram metafor-teorien.

### **Et tilstandsforhold**

Fram til nå i analysen har jeg blant annet sett etter bevegelsesinntrykk, både på grunn av deres forbindelse til handling-lyd og på grunn av deres forbindelse til romlighetsmetaforer og eksperimentelt meningspotensiale. I teorikapitlet nevnte jeg hvordan *image schemata* (Johnson 1987) var en mental romlig forståelse skapt av tilbakevendende mønstre i vår daglige erfaring med verden. Device Connect-lyden demonstrerer dette prinsippet på en god måte ved å spille på symmetri mellom ulike kontrasterende forhold.

Handlingen som hører til lydhendelsen Device Connect går ut på at en enhet som tidligere ikke var til stede i datamaskinens virtuelle verden, nå er blitt gjenkjent og er til stede. Dette som en konsekvens av at en handling har funnet sted i brukerverdenen. Et eksempel kan være at en bruker har koblet en kabel inn i datamaskinen. En slik tilkobling fører med seg en armbevegelse, men den kan også sees som et kontrastforhold hvor en gjenstand som tidligere ikke hørte til noe, nå hører til noe fordi gjenstanden er koblet til datamaskinen. Denne



tilhørigheten<sup>13</sup> kan vi enkelt se i brukerverdenen gjennom fysiske kabelforbindelser, men nå som trådløs tilkobling blir mer og mer populært, minnes vi på at det også finnes slike tilhørighetsforhold i datamaskinens virtuelle verden.

Kontrasten mellom at noe hører til eller ikke hører til noe annet, kan sees som en romlighetsmetafor. Hvis vi tar utgangspunkt i at tilhørighet er positivt, mens å stå utenfor er negativt, vil den kjente ‘Glad er opp, trist er ned’-romlighetsmetaforen føre til at tilhørighet tilsvarer ‘opp’ mens å stå utenfor er ‘ned’. Dermed får vi en emosjonell akse som både kan beskrives gjennom bevegelsen ut-inn og ned-opp. Ut-inn-bevegelsen skapes av brukeren, ned-opp gjenspeiles i intervallet mellom de den høye og den lave tonen til systemlydene i Device Connect. To forhold ved denne lydhendelsen vil komplisere resonnementet, og det er forsinkelse og virtuelle handlinger. De kommer jeg tilbake til i den multimodale analysen min.

Symmetrien mellom den tilkoblende bevegelsen i brukerverdenen er noe som vi kan oppfatte gjennom eksperimentelt meningspotensiale, i den forstand at vi forstår forholdet mellom de to tonene i systemlyden ut fra vår kroppslige forståelse, ved å koble en gjenstand til en annen. Betydningen til systemlydene blir dermed ‘tilkobling’, og fordi bevegelsen går oppover, blir den en romlighetsmetafor for ‘vellykket’, altså betyr systemlydene ‘vellykket tilkobling’. Jeg nevnte da jeg analyserte typologien til Windows XP-lyden at den lange halen forlenget bevegelsesinntrykket, og i dette tilfellet betyr det at bevegelsen først stiger og så blir hengende. Ting som henger i friluft gjør det stort sett med god grunn, og jeg tolker denne forlengede bevegelsen som et tegn på forventning, at systemlyden forventer at vi skal gjøre noe med den enheten som er gjenkjent av systemet. Begge systemlydene kan dermed bety en vellykket tilkobling, men i Windows XP har systemlyden dessuten et innslag av forventning i seg.

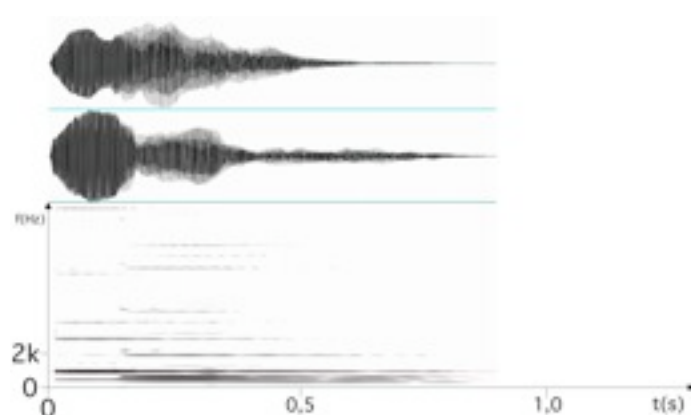
Avslutningsvis i dette resonnementet bør jeg nevne at Device Connect har to søsterhendelser: ‘Device Disconnect’ og ‘Device Failed To Connect’. Den første av de to har speilvendt tonemønster fra Device Connect, og følgelig blir da betydningen motsatt, nemlig en frakobling. Her går bevegelsen nedover, og dette kan gi to mulige tolkninger av den emosjonelle betydningen: enten er det negativt at enheten ikke lenger ble gjenkjent av datamaskinen (på grunn av strømbrudd eller andre feil), eller bevegelsen nedover kan sees som en lettelse for at vi var ferdige med det vi skulle gjøre, på samme måte som skulderene våre

---

<sup>13</sup> Tilhørighet, eller *containment*, er en av Johnsons kategorier av *image schemata* (Johnson 1987)

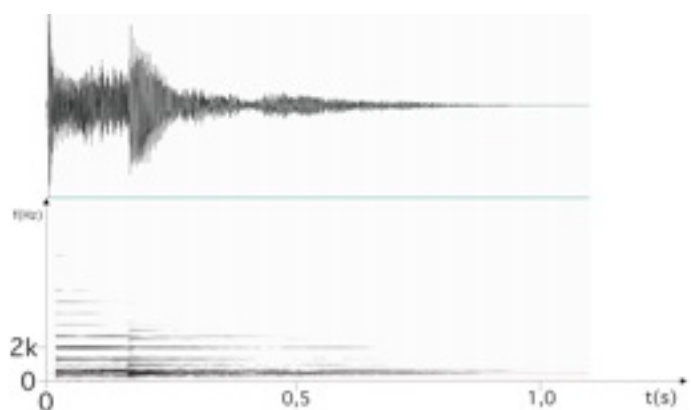
senker seg når vi trekker et lettelsens sukk. Device Failed To Connect består av tre anslag av den første tonen (den lave), og kan bety at bevegelsen står fast på første trinn. Denne betydningen er helt klart lettere å forstå hvis vi allerede har erfaring med de to andre lydene, ellers ville tolkningen sannsynligvis bli ganske lik noe slikt som vi hadde i Battery Critical for Windows XP, hvor bevegelsen var en hamring eller bremsing. I den neste lydhendelsen jeg går gjennom, får vi se lignende virkemidler som de jeg nettopp har vært gjennom.

## Hendelse 6 av 6: New Mail



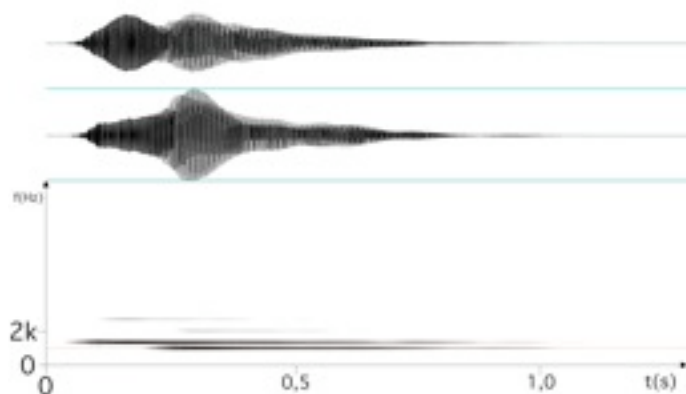
Figur 3m: New Mail (Windows 2000)

**Format:** 16-bit PCM 22,05 kHz Stereo  
**Kraftigste amplitude:** -14,4 dBFS  
**Varighet:** 1,3 s.  
 (Anslag 0,02 s. Kropp 0,5 s. Hale 0,8 s.)  
**Toneinnhold:** A som går en kvint ned til D



Figur 3n: New Mail (Windows XP)

**Format:** 16-bit PCM 22,05 kHz Mono  
**Kraftigste amplitude:** -18,7 dBFS  
**Varighet:** 1,1 s.  
 (Anslag 0,005 s. Kropp 0,2 s. Hale 0,9 s.)  
**Toneinnhold:** Diss som går en kvart ned til Aiss



**Format:** 16-bit PCM 44,1 kHz Stereo

**Kraftigste amplitude:** -22,5 dBFS

**Varighet:** 1,3 s.

(Anslag 0,15 s. Kropp 0,3 s. Hale 0,85 s.)

**Toneinnhold:** Diss som går en kvart ned til Aiss

Figur 3o: New Mail (Windows Vista)

Disse tre systemlydene for New Mail er rimelig like, så jeg kommer ikke til å gå i dybden på de typologiske og morfologiske kvalitetene. Lydene kjennetegnes av to anslag med to ulike hørbare toner, på samme måte som ved Device Connect. Intervalltrinnet er større i Windows 2000 i forhold til de andre systemlydene, men retningen er nedover i alle tilfeller. Det er akkurat de samme hørbare tonene som brukes både i Windows XP og Vista, så da er det grunn til å spørre seg om årsaken til det ligger i at konvensjonene for denne lydhendelsen har fått etablere seg. Dette kommer jeg til å ta opp når jeg oppsummerer de historiske trekkene jeg har funnet i alle lydhendelsene.

Alle systemlydene har tonalt sentrum høyt oppe, og lydobjektene har derfor liten masse, kanskje for å ligne den lette tyngden til et papirark. Anslaget til Windows XP er langt kortere enn hos de to andre systemlydene, og preges derfor mer av et slag eller treff enn de andre lydene, som har såpass langt anslag, kropp og hale at de er mer utholdte enn impulsive. En utholdt typologi i Windows 2000 og Vista kan tyde på en mer flyvende bevegelse, som et kast eller noe streifende. Det er også en del energiforflytning mellom kanalene i disse lydene, hvor bevegelsen veksler fra side til side på en måte som ligner måten et papirark langsomt svever fra side til side når det er i fritt fall. Lyden til Windows XP er i mono, og har derfor ingen slike sidelengsbevegelser, men vi hører til gjengjeld mer av rommet som lydobjektet befinner seg i, takket være romklngen.

Betydningen som kommer av typologiske og morfologiske trekk kombinert med bevegelse, gjenspeiler for alle systemlydene at vi her har med en flyvende liten lett gjenstand å gjøre, og i

e-postsammenheng er det naturlig å tro at gjenstanden er et brev. I Windows 2000 har vi et brev som lander mens det svever fra side til side. I Windows XP hører vi et brev som treffer noe i et lukket rom, kanskje de harde kantene til konvolutten treffer kantene til en overdrevent stor postkasse som lydobjektet ramler inn i? Av og til kan det være nødvendig å overdrive virkemidlenes størrelse for å få fram riktig betydning, så i dette tilfellet kan det hende at vi ikke bryr oss om at den romstørrelsen vi hører er overdrevent stor. I Windows Vista ligner klangen i lydobjektet på lyden når vi plystrer, og toneintervallet ligner det en del norske hundeeiere plystrer når de kaller tilbake hunden sin. Dette er en tydelig kulturelt avhengig tolkning, men den er sterk nok til at jeg legger til ‘anrop’ sammen med de to andre betydningene som jeg hadde funnet i systemlyden fra før. For alle tre lydene eksisterer det et par andre tilsvarende konnotasjoner som kan passe for noen lesere, for eksempel lyden av en heis som når riktig etasje, eller lyden som kommer over høyttaleranlegget før stasjonspersonalet melder at toget ditt ankommer en platform. ‘Ankomst’ blir derfor enda en mulig betydning som kan trekkes ut av dette tonemønsteret.

Jeg vil hevde at vi i likhet med Device Connect her har store deler av betydningen liggende i intervallet mellom lydens to hørbare toner. I denne lydhendelsen stiller imidlertid de multimodale forbindelsene seg noe annerledes. For Device Connect utspant det seg en lydskapende handling både i brukerverdenen og i datamaskinens virtuelle verden. Her er det, så lenge vi ser bort fra unntak som at en bruker klikker på ‘hent ny epost’, bare en handling i den virtuelle verdenen. Dette anser jeg likevel ikke som noen reduksjon i betydning, siden jeg tror vi i Device Connect er fornøyd med den betydningen som vi finner først, gitt at handling-lyd-slektskapet virker logisk. Uansett om vi skulle se ‘færre multimodale bevegelsesrelasjoner’ som en reduksjon i betydning eller ikke, hjelper det på meningsformidlingen at morfologien for systemlydene i denne lydhendelsen stemmer overens med vår oppfatning av et brev.

Lydstyrken blir svakere for hver systemutgave i denne lydhendelsen, og fordi lydene ellers er såpass like, er det rimelig å tro at systemlydene til de nyere systemutgavene oppfattes som mildere enn i eldre systemutgaver. Det ble den siste kommentaren i min deskriptive analyse, så nå passer det å oppsummere og se hvilke historiske trekk jeg kan komme fram til av mine data.

### 3.2.1 Oppsummerende tabell, deskriptiv semiotisk analyse

| Hendelse                 | Typologi          | Morfologi  | Bevegelse                       | Semiotisk tegn                       |
|--------------------------|-------------------|--|---------------------------------|--------------------------------------|
| Critical Stop (2000)     | impulsiv          | liten og lett gjenstand i stort rom, hardt og glatt treverk      | slag / treff                    | skarphet, dramatikk                  |
| Critical Stop (XP)       | impulsiv/utholdt  | stor og tung påtrengende klump, myk og knudrete                  | bremsing / dunking              | obstruksjon, dramatikk               |
| Critical Stop (Vista)    | impulsiv          | stor og mellomtung plate eller vegg av metall eller glass        | slag / treff                    | skarphet, dramatikk                  |
| Default Beep (2000)      | impulsiv/utholdt  | stor, mellomtung, hard og glatt gjenstand plassert mellomhøyt    | slag / vifing                   | anonym oppmerksomhet                 |
| Default Beep (XP)        | impulsiv          | mellomstort og mellomtungt treverk, lavt plassert                | slag / dunking                  | anonym oppmerksomhet                 |
| Default Beep (Vista)     | impulsiv          | mellomstort og lett glass, høyt plassert                         | slag / knips                    | anonym oppmerksomhet                 |
| Battery Critical (2000)  | impulsiv          | stor, mellomtung, hard og glatt gjenstand plassert i midtsjiktet | slag / vifing                   | anonym oppmerksomhet                 |
| Battery Critical (XP)    | impulsiv/iterativ | stor og tung påtrengende klump, myk og knudrete                  | hamring / bremsing              | tapt levedyktighet                   |
| Battery Critical (Vista) | utholdt           | liten og lett, glassaktig, høyt plassert                         | løfting                         | resignasjon                          |
| Notification (2000)      | impulsiv          | liten og lett gjenstand i stort rom, hardt og glatt treverk      | slag / dask                     | skarphet, åpenhet, tilbakemelding    |
| Notification (XP)        | impulsiv          | liten og lett elastisk boble, tett inntil                        | oppsvulming, deretter sprekking | plutselig bemerkning, forventning    |
| Notification (Vista)     | impulsiv          | stor og tung glatt metallgjenstand, tett inntil                  | forbipassering / streifing      | bortgjemt bemerkning                 |
| Device Connect (2000)    | ---               | ---  | ---                             | ---                                  |
| Device Connect (XP)      | impulsiv          | hardt treverk. lite rom  | tilkobling                      | vellykket tilkobling, forventning    |
| Device Connect (Vista)   | impulsiv          | elastisk gummi. lite rom   | tilkobling                      | vellykket tilkobling                 |
| New Mail (2000)          | impulsiv/utholdt  | liten og lett gjenstand  | landing / sveving               | lettelse, kunngjøring, ankommet brev |
| New Mail (XP)            | impulsiv          | liten og lett gjenstand i stort rom                              | ramlende treff                  | lettelse, kunngjøring, ankommet brev |
| New Mail (Vista)         | utholdt           | liten og lett gjenstand, med klang lik plystring                 | kast / flyvende                 | lettelse, anrop, ankommet brev       |

Tabell 3a: Egenskaper ved lydhendelsene

### 3.3 Problemstilling 1: Historisk utvikling

Nå som vi har sett på lydene i detalj, passer det å se dem i forhold til den første problemstillingen min, som dreier seg om hvordan de semiotiske trekkene til systemlydene i Microsoft Windows har utviklet seg de siste årene. Funnene som vi nettopp har gått gjennom, tyder på at det er mange motstridende virkemidler på tvers av systemversjonene, og kombinert med tegnets polysemiske natur ville det nok vært vanskelig å få noe særlig resultater av å foreta en ‘nedenfra og opp’-analyse ([Godøy 1997:170](#)) av dette materialet. Derfor skal jeg bruke de neste avsnittene på å presentere en del historiske trekk som jeg kan utlede fra det vi nettopp har vært gjennom. Jeg begynner med se på utviklingen i de rent tekniske forhold ved lydene. Deretter ser jeg på den historiske utvikling i typologi og morfologi, toneinnhold og til slutt utvikling i systemlydenes betydning.

#### **Tekniske forhold**

Det tekniske nivået til systemlydene har utviklet seg gjennom de ulike systemversjonene, selv om det ikke ser ut til å skyldes noen overveiende teknologisk determinisme. Med det mener jeg at selv om datakraften øker etterhvert som vi får nyere og raskere datamaskiner, ser det ikke ut til at avspillingsforholdene til systemlydene stiller noen strengere krav til datakraft. Faktisk går vi fra stereolyd i Windows 2000 og tilbake til mono i Windows XP, for så å gå tilbake til stereo i Windows Vista. Noe av dette får en konsekvens for meningsformidlingen, da systemlydene i Windows XP ikke kan bruke panorering mellom kanalene som et virkemiddel til å beskrive lydobjektens morfologi. Det er ikke sikkert at det gjør så mye, siden et godt stereobilde vil være avhengig av at brukeren også befinner seg i riktig posisjon i forhold til høyttalerne til enhver tid, og dermed er det risikabelt hvis viktig betydning skulle ligge i stereobildet alene. Når det gjelder lydhendelser, er de imidlertid høyere i de to senere systemversjonene, og samplingsfrekvensen til lydene har gått opp fra 22,05 til 44,1kHz, så en viss teknisk tilpasning har likevel funnet sted. Inndelingen av lydhendelser og tilegning av systemlyder i standardoppsettet har holdt seg stabilt gjennom systemene, så det virker ikke som om Microsoft har brutt noen konvensjoner med vilje.

På det rent lydstyrkemessige har jeg sett følgende utvikling: Windows 2000 var noenlunde jevnt i lydstyrke på tvers av systemlydene. Jeg har analysert i alt tre lydfiler derfra, hvor to av dem har samme lydstyrke, og den tredje (den med ‘skarp’ betydning) er mindre enn 6 desibel

sterkere. Windows XP hadde samme gjennomsnittlige lydstyrke i systemlydene som Windows 2000, men langt større forskjell i lydstyrke mellom hver enkelt systemlyd. På det mest ekstreme er det over 18 desibel forskjell fra en lyd til en annen. Det tilsvarer en dobling av effekt tre ganger, og da kan en lyd teknisk sett være åtte ganger kraftigere enn en annen. Med slike forhold kan man risikere at noen lyder forsvinner i lydbildet fordi de er for svake til å bli hørt og tolket, eller at andre lyder blir spilt altfor sterkt, og på den måten blir plagsomme. At det eksisterer så store forskjeller i lydens dynamikk, står i sterk kontrast til nyere trender innenfor radio- og fjernlynslyd ([Maaso 2002](#)) hvor lydens dynamikk ofres til fordel for at lyden skal stikke seg ut i mengden. Tydeligvis er datamaskiner et ganske annet medium enn radio og TV i denne sammenheng. Dynamikk kan brukes til å forsterke de følelsesmessige reaksjonene som lyden kan bringe med seg ([Van Leeuwen 1999:161](#)), og det er mulig at det var det Microsoft prøvde på i Windows XP. I Windows Vista har de imidlertid redusert disse forskjellene igjen. Her er det gjennomsnittlige lydnivået til systemlydene sunket med 6 desibel i forhold til de to foregående systemene. Den individuelle forskjellen i lydnivå mellom systemlydene er samtidig redusert til en forskjell på 9 desibel. Dette tyder på at lydene generelt er blitt svakere, men at de svakeste lydene ikke står i like stor fare for å drukne i resten av lydmiljøet som det de sto i fare for å gjøre i Windows XP.

### **Typologi og morfologi**

Når det gjelder systemlydenes klangfarge og omhyllingskurve, er det relativt store likheter innenfor hvert enkelt system. Generelt blir det mer energi i diskanten etter hvert som systemene blir nyere. Windows 2000 har en tett frekvensansamling, Windows XP noe mer spredt, mens Windows Vista har frekvenser spredt såpass bredt at det av og til er vanskelig å skille enkelttoner fra klang. Omhyllingskurven til Windows 2000-lyder har typisk et kort anslag og en lang hale, Windows XP har kort anslag, lang kropp og kort hale, mens Windows Vista har et lengre anslag og lengre hale enn de to foregående systemene. Systemlydene til Windows Vista får derfor et mykere preg enn 2000 og XP, fordi lydstyrken hverken kommer eller forsvinner like brått som i de to eldre systemene. Jeg finner også eksempler på at de lydene som har vært skarpe og med stor masse, blir forskjønnet i senere systemer, og da særlig i Windows Vista. Forskjønnelsen innebærer en omhyllingskurve med lengre anslag for å redusere skarphet, lavere lydstyrke, og et fravær av kraftig bassinformasjon slik at lydobjektene får mindre masse og dermed ikke blir like dominerende. Det sterkeste eksempelet

på dette inntreffer i lydhendelsen Battery Critical, der det er vanskelig å komme fram til en betydning som virker troverdig i forhold til lydhendelsen.

Forbindelsen mellom typologi, morfologi og den semiotiske betydningen som jeg kom fram til, var avhengig av at jeg først kom fram til en bevegelse før jeg klarte å utlede noen betydning fra funnene mine. Dette betyr at selv om jeg klarte å finne en brukbar mengde med morfologiske trekk ved en systemlyd, var det ingen direkte forbindelse mellom morfologi og semiotiske tegn. Det bevegelsesinntrykket som var nødvendig for å finne betydning, oppsto som en kombinasjon av en grov bevegelse jeg fikk fra de typologiske kvalitetene og en nyanse av denne bevegelsen som jeg fikk fra de morfologiske kvalitetene. Dette kan tyde på at vi i handling-lydslektskap først grovsorterer handlinger som vi hører, og deretter kombinerer dette inntrykket med lydens morfologi til å finsortere og danne oss en mening om lydens tilhørighet og betydning.

## **Toneinnhold**

Tonaliteten til systemlydene ser ut til å være uniform innenfor hver enkelt systemversjon, med lyder som tar utgangspunkt i tonen G i Windows 2000, C i Windows XP og A i Windows Vista. Imidlertid er dette tre forskjellige tonaliteter, så i den grad vi skal kunne se likheter i toneinnhold, må disse likhetene komme i andre egenskaper som for eksempel tonehøyde og akkorder. Det er få eller ingen mønstre i tonehøyde, hverken hvis vi hører lydene kronologisk eller innenfor ett enkelt system. Når det gjelder akkorder og intervaller, er det bare noen få mønstre på tvers av systemene. Disse merker vi tydeligst i det jeg kaller *intervallbaserte systemlyder*, som Device Connect og New Mail. I førstnevnte har vi samme intervall, men i to forskjellige tonearter. I sistnevnte har vi en forskjell i toneart og intervall mellom Windows 2000 og XP, men så beholdes akkurat de samme tonene i Windows XP og Vista. Tonene beholdes, til tross for at det går på bekostning av toneforholdet opp mot de andre Vista-lydene. Jeg tolker dette som en indikasjon på at Microsoft var så fornøyd med New Mail-lyden i Windows XP at de ikke så det som nødvendig å endre noe annet enn overfladiske kvaliteter i denne systemlyden til Windows Vista.

## **Betydning**

Det var mulig å tolke betydninger ut av alle systemlydene, hvor betydningen i større eller mindre grad stemte overens med den virtuelle handlingen som lyden skulle representere.



Riktignok ble antagelig den betydningen jeg fant noe farget av at jeg analyserte systemlydene vel vitende om hvilken hendelse de hørte til. På den annen side var det så stort sprik i resultatene mine til tross for at jeg analyserte ovenfra og ned, at det er tvilsomt om en ‘nedenfra og opp’-analyse ville vært særlig hensiktsmessig med så mange ukjente variabler involvert.

Et spesielt funn lå it at Windows 2000 brukte samme lyd for to hendelser som betydningsmessig lå langt fra hverandre: Critical Stop og Notification. Selv om lydenes betydning både spiller på skarphet og åpenhet, vet brukeren ikke hvilken av hendelsene det gjelder bare ved å høre lyden alene. Derfor vet ikke brukeren ved å høre lyden, om datamaskinen krever umiddelbart tilsyn eller om det kan vente.

De lydene som hadde mest uklar betydning, var Notification i Windows 2000, Notification i Vista og Battery Critical i Vista. Uklarheten skyldes at bevegelsesinntrykket og morfologien jobbet i forskjellige retninger, og ingen av disse retningene hadde noen særlig tilknytning til den handlingen som lyden hørte til. Lydene som var lettest å finne betydning i var Critical Stop i Windows XP, Device Connect og New Mail (uavhengig av operativsystem). Critical Stop var lett å se betydning i fordi bevegelsesinntrykket og morfologien jobbet i samme retning. Device Connect og New Mail hadde bevegelsesinntrykk som stemte på tvers av de multimodale forbindelsene, og for dem virket det som om konvensjonene var falt på plass.

## **Oppsummering**

For å oppsummere funnene så langt, kan den historiske utviklingen virke relativt kaotisk og lite konsekvent når vi hører på en lydhendelse kronologisk gjennom de ulike systemene. Unntaket ligger i klangen og omhyllingskurven til systemlydene, hvor nyere systemer får flere overtoner og mykere anslag, og enkelte lyder blir forskjønnet. Mykere anslag, mindre bassinformasjon og lavere lydstyrke. Av og til gjøres dette selv i situasjoner hvor slike endring går på bekostning av meningsformidlingen. Likevel er forholdene langt mindre kaotisk innenfor hvert enkelt operativsystem, hvor både tonalitet og klangfarge holder seg noenlunde likt på tvers av alle lydhendelsene. Sannsynligvis sørger denne interne likheten for å skape en viss harmonisk orden når vi hører lydene i vår daglige databruk, som igjen kan tydeliggjøre at vi betjener et medieprodukt, nemlig Microsoft Windows. Dette vil jeg komme tilbake til i siste del av denne oppgaven, når jeg drøfter hva Microsoft prøver å få til med lyddesignet sitt.

Som vi har sett, går det an å utlede en semiotisk betydning fra alle lydene, hvor denne betydningen står i samsvar med den respektive lydhendelsen. Imidlertid blir det spekulativt å hevde at lydene betyr det samme som lydhendelsen. Dette skyldes naturligvis tegnets polysemiske natur, men også den store variasjonen i oppmerksomhet og lydforståelse hos forskjellige brukere, samt at avspillingsutstyr og situasjonen som lydhendelsen dukker opp i, vil påvirke forståelsen av lyden i forhold til konteksten. I det minste kan jeg hevde at betydningen av en systemlyd virker fornuftig når den kommer sammen med lydhendelsen, gitt at vi på forhånd (av erfaring) vet hva slags hendelse som systemlyden representerer. Her foregår det altså en meningsutveksling mellom lyden og lydhendelsen, men vi må ikke glemme at vi kan få betydninger fra andre samtidige modi som skrift og visuelle tegn. Neste punkt på programmet blir å se hvordan lyden forholder seg til disse andre medieinntrykkene, og på hvilken måte betydningen kommer fram.

## 4 Multimodal analyse

---

**H**ittil har jeg sett på systemlydenes betydning ut fra lydenes interne egenskaper. Selv om det var mulig for meg å utlede en betydning ut fra lyden alene, bruker vi langt fra bare lydinformasjon til å danne vår formening om betydningen i en lydhendelse. Det er mange andre samtidige inntrykk som spiller inn, deriblant visuelle og kroppslige inntrykk. Mønsteret mellom de ulike betydningene sier noe om systemlydenes betydningsstrøm, og denne betydningsstrømmen skal jeg nå studere nærmere.

I [Kapittel 2.3](#) tok jeg opp hvordan tegnet kunne få ny betydning gjennom fenomener som konnotasjon, forankring og avløsning, og hvordan dette kunne forsterkes multimodalt gjennom provenance og eksperimentelt meningspotensiale. Samtidig ser handling-lyd-slektskapet på en annen betydningsstrøm, nemlig hvordan lyd skapes av et objekt i bevegelse og hvor bevegelsesinntrykkene eksisterer på flere samtidige plan: kroppslige, mentale, fysiske og virtuelle. Her vil jeg prøve å kombinere disse ulike synene i en multimodal analyse av noen konkrete situasjoner på datamaskinen hvor systemlyder inntreffer.

### 4.1 Sannhetsforsterkning

Som jeg nevnte i teorikapittelet, har Roland Barthes' teorier om forankring og avløsning høstet kritikk i nyere forskning, blant annet fordi termene forutsetter et styringsforhold hvor en modalitet er underlagt en annen. Kritikerne etterlyser en bedre forståelse av hvordan ulike modaliteter gir hverandre merverdi. For å imøtekomme dette, vil jeg i denne analysen studere hvordan lyd skaper merverdi gjennom sannhetsforsterkning. Ifølge [Theo Van Leeuwen](#) (1999:160-161) finnes det som sagt fire sannhetskategorier: naturalistisk, abstrakt, skjematisk og sensorisk. Jeg tenker å varme opp denne teorien før analysen min, ved å vise hvordan den kan anvendes på et velegnet konstruert eksempel.

Tenk deg at du hører på et gammelt opptak av en klassisk konsert. Grunnen til at du akkurat nå hører på et opptak, og ikke en konsert direkte, er fordi du da lytter til objekter i en virtuell verden, på lignende måte som lyder som kommer fra et brukergrensesnitt. Det virker troverdig at det du hører gjenspeiler et konsertopptak, fordi du gjenkjenner de ulike klassiske

instrumentene. Samtidig er det litt skurring i opptaket. Skurringen øker den *naturalistiske* sannheten, fordi den får deg til i enda større grad akseptere at dette var et opptak gjort for mange år siden. Ingen skurring i opptaket gir mindre naturalistisk sannhet. For mye skurring eller støy av en type som ikke er troverdig, ville derimot brikke over i det surrealistiske, så det er viktig å være klar over at sannhetsforsterkning ikke foregår innenfor en skala med absolutte sluttverdier.

I denne konserten du lytter til, beveger tonene til instrumentene seg langsomt nedover, for å symbolisere sorg. Etter dette blir lydene svakere og langsommere, for å symbolisere død. Dette er virkemidler som er godt etablert innenfor klassisk musikk, for eksempel ifølge Barokkens affektlære (Leman 2007). Midt i disse konturene spiller en trompet en veldig hurtig melodi som beveger seg freidig opp og ned, og den havner i kontrast til de andre musikalske inntrykkene. Du syns kombinasjonen av det komiske og tragiske blir et uttrykk for likegyldighet til døden, og opplever en høy *abstrakt* sannhet. Et eksempel på musikk med med lav abstrakt sannhet kan være en annen innspilling av samme musikkstykke, hvor du fremdeles hører tonene som går nedover, men tempoet og lydnivået holdes jevnt og skaper ingen kontrast til trompetmelodiens tempo.

Underveis i musikkstykket merker du at instrumentene blir spilt i et repeterende tidsmønster, og dette tolker du om til et visst rytmisk tempo. Trommene forsterker og tydeliggjør dette tempoet fordi de peker seg ut som mønstre i intensitet som passer med resten av instrumentene. Tonene danner et mønster som du tolker om til toneart, og du hører at den skifter underveis i et senere parti av konsertopptaket. Dette er en *skjematisk* sannhetsforsterkning som tydeliggjør mønstre som kan omsettes til konkrete verdier, i dette tilfellet tempo og toneart. Et eksempel med lav skjematisk sannhetsforsterkning kan komme senere i musikkstykket når det bare spilles foliner som glir over i hverandre rytmisk og hvor tonene ikke stemmer med hverandre. I det ene tilfellet orden, i det andre kaos.

Noen dager senere er du innom nærmeste bensinstasjon og kjøper en musikk-DVD med nettopp samme konsertopptak. Du lot deg friste av ordene 'Digitally remastered full surround sound' på forsiden. Opprømt ringer du vennen din som har hjemmekino, og dere møtes og lytter til konsertopptaket sammen. Skurringen i opptaket er fjernet, romklang er lagt til, og lydkilder er separert til forskjellige lydkanaler. Det føles som du sitter midt i orkesteret. Du har

nettopp vært vitne til en *sensorisk* sannhetsforsterkning. Til forskjell fra den naturalistiske sannhetsforsterkningen, har lydens attributter ikke bidratt til mer troverdighet om at du lytter til et konsertopptak. I stedet bruker de bevisste trekk for å fremkalle en følelsesmessig reaksjon. I dette tilfellet kan din følelsesmessige reaksjon ha vært bakoversveis og tårer i øynene, men den kan også ha vært frykt i de situasjonene hvor lyden ble spilt så høyt at energien i bassen fikk deg til å skvette.

## 4.2 Multimodal analyse

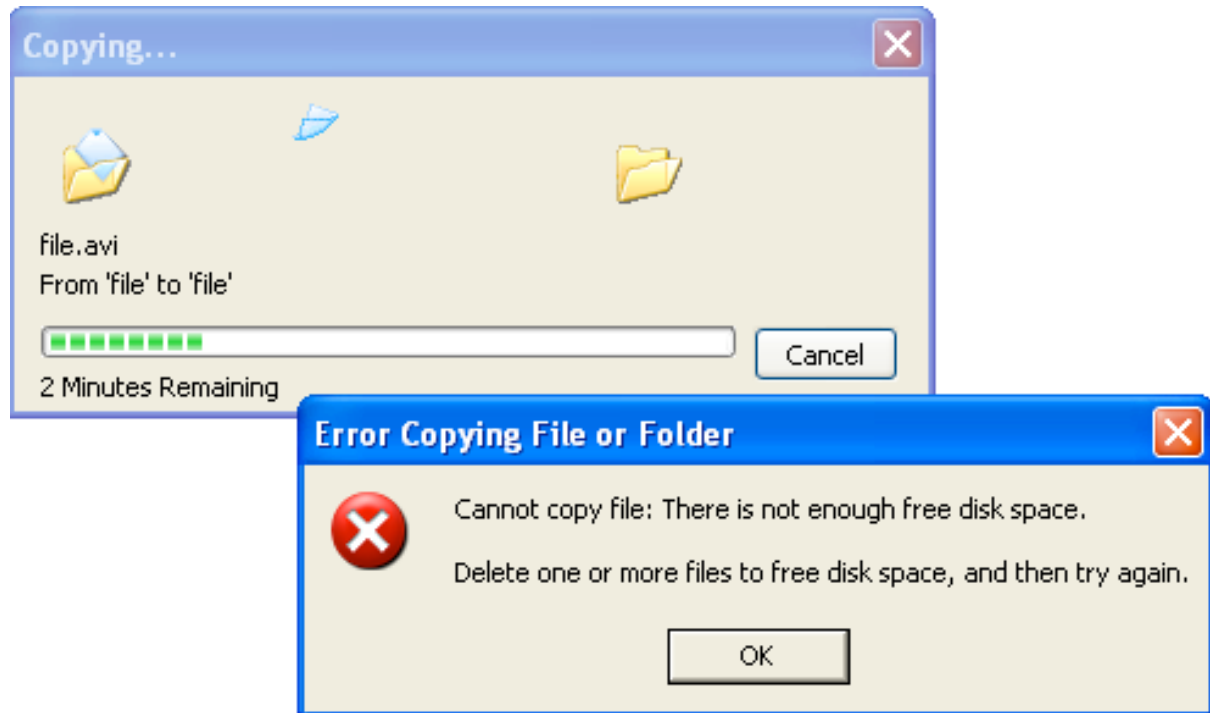
Nå som vi har et inntrykk av hvordan vi kan oppdage ulike former for sannhetsforsterkning, vil jeg kombinere denne teorien sammen med å se etter forankring, avløsning og konnotasjon i analyseobjektene. På den måten vil jeg både se på hvordan betydning skapes mellom ulike modaliteter, om det skapes nye tegn gjennom konnotasjon og om lyden har en merverdi i denne sammenhengen.

I resten av dette kapittelet vil jeg se på betydningsstrømmen i fire konkrete lydhendelser. De tre første er fra Windows XP: Critical Stop, Notification og Device Connect. Den siste er fra Windows Vista: Battery Critical. Jeg har valgt akkurat disse lydhendelsene av to årsaker. For det første har jeg sett etter situasjoner hvor det er tydelige mønstre mellom lydens betydning, skrift og andre visuelle inntrykk. For det andre har jeg sett etter situasjoner hvor lyden har ulike former for sannhetsforsterkning. I motsetning til den semiotiske analysen jeg gjorde tidligere, er jeg ikke på jakt etter noen historisk utvikling. Likevel tror jeg at funnene mine kan gi en indikasjon på sammenhenger som også kan gjelde de systemlydene som jeg ikke har analysert multimodalt.

### Hendelse 1 av 4: Critical Stop (Windows XP)

For å støte borti denne lydhendelsen har vi satt igang en kopiering av noen filer, og da dukker det opp et vindu som grafisk viser et ark som kastes fra en mappe til en annen. Under dette står det hva som kopieres, hvor det kopieres fra og hvor det kopieres til. Vi kan holde øye med prosessen via en progresjonslinje som langsomt fylles opp og skrift som sier hvor mye tid som gjenstår. Plutselig går maskinen tom for plass på harddisken, og vi får opp en feilmelding som i *Figur 4a*. For å analysere denne situasjonen multimodalt, begynner jeg med å redegjøre for noen visuelle tegn som finnes her, og hvordan betydningen kommer fram gjennom disse

tegnene. Deretter skal jeg vurdere hvor stor sannhetsforsterkning som kommer fram i systemlydens modalitet, og hvordan denne sannhetsforsterkningen påvirker meningsformidlingen.



Figur 4a: Skjerm bilde fra Critical Stop (*Microsoft 2001*)

## Visuelle tegn

I denne feilmeldingen eksisterer det en rekke visuelle tegn, presentert som skrift og grafiske ikoner. Skriften forteller oss over tre linjer navnet på feilen, deretter hva som er feilen, og til slutt en oppfordring til hva vi bør gjøre. Helt til slutt kommer bokstavene 'OK'. Disse siste bokstavene har en funksjon i brukergrensesnittet, da de er omringet av noen linjer som vi gjennom brukergrensesnittets konvensjoner forstår at utgjør en trykkbar knapp. 'OK' blir dermed *signifikanten* i dette tegnet som knappen utgjør, og *signifikatet* blir ideen om at knapper er noe som kan trykkes på i brukergrensesnittet.

Det er ikke bare 'OK'-knappen som fungerer som et grafisk tegn i denne feilmeldingen. To steder har vi en hvit 'X' på rød bakgrunn. Den ene er en rød sirkelbakgrunn, mens den andre befinner seg oppe til høyre i feilmeldingen og har en kvadratisk rød bakgrunn. Selv om disse grafiske ikonene er såpass like, har de forskjellig betydning i brukergrensesnittet. Den røde sirkelen til venstre er et grafisk symbol som hører til denne konkrete feilmeldingstypen, på

samme måte som lyden for Critical Stop også hører til denne feilmeldingstypen. Den røde firkanten med hvit X som er oppe til høyre i vinduet er der ikke fordi dette er en feilmelding, men fordi det er et vindu. Vi vet gjennom vår erfaring med brukergrensesnittet Windows at en X oppe i høyre hjørne av et vindu er en lukkerute, en knapp som lukker vinduet. Vi kan altså fjerne denne feilmeldingen på to måter: enten kan vi lese tegnet for at vi har lest meldingen og akseptert den ('OK'), eller vi kan lese tegnet på at dette er et vindu som lar seg lukke. Resultatet blir det samme.

Hvis vi derimot prøver å trykke på bildet av en hvit X på rød sirkel som befinner seg til venstre i vinduet, skjer det ingenting. Dette grafiske ikonet har ingen funksjon i brukergrensesnittet annet enn å bære informasjon om handlingen som har skjedd i den virtuelle verdenen.

Rødfargen kan en bruker med normalt fargesyn tolke som et tegn på fare, men også som et tegn på aggresjon ut fra at vi mennesker kan bli 'røde av sinne'. En annen tolkning er at rød er fargen for en obstruksjon eller et forbud, på samme måte som et rødt trafikklys betyr 'stopp'. Det hvite krysset blir enda et tegn som støtter opp om denne tolkningen.

En klassisk tolkning av Roland Barthes innebærer at skriften her *forankrer* bildet ved å fortelle oss at datamaskinen ikke har klart å fullføre handlingen som vi ba den om å utføre. Det grafiske ikonets betydning 'forbud' velges vekk til fordel for betydningen 'stopp'. Samtidig *avløser* skriften bildet ved å gi oss detaljer om hvordan denne handlingen ble avbrutt, i dette tilfelle at harddisken er full. Dermed kan den røde sirkelen med X få en betydning lik 'diskproblem' eller tilsvarende.

Problemet med tolkningen av betydningsstrømmen ovenfor er at den ikke tar hensyn til at erfarne databrukere leser feilmeldingen på en annen måte. Når feilmeldingen inntreffer, kombineres handlingen i bakgrunnen med grafiske ikoner som viser at dette er en 'stopp-feilmelding'. Det eneste som mangler av informasjon før brukeren kan trykke 'OK', er årsaken til hvorfor feilmeldingen oppsto. Øyet bruker betydningen fra de andre visuelle inntrykkene til å velge ut en bestemt betydning fra skriften, som i dette tilfellet er 'not enough free disk space'. De andre linjene med skrift velges bort fordi de grafiske ikonene *forankrer* en spesiell betydning i skriften. Førstegangs databrukere vil kanskje ikke lese feilmeldingen på denne måten, men konvensjonene er såpass innarbeidet i datasystemer at det ikke skal leses mange feilmeldinger før en slik hurtiglesing er innarbeidet.

## Lydens sannhetsforsterkning

Grunnen til at jeg så langt har brukt plass på å gå gjennom den visuelle betydningsstrømmen i feilmeldingen, er fordi vi i denne lydhendelsen er vitne til at lyden står i et lignende forhold til feilmeldingen som det grafiske ikonet av den røde sirkelen med hvitt kryss. For å illustrere dette, vil jeg studere hvordan lydens modalitet kommer fram gjennom de fire formene av sannhetsforsterkning, før jeg ser hvordan denne betydningsstrømmen forholder seg til andre samtidige inntrykk.

- Naturalistisk?

Det er ingen sterk naturalistisk sannhet til stede, men bare i den forstand at lyden gir et visst bremseinntrykk. Om dette bremseinntrykket virker troverdig i forhold til handlingen, vil være avhengig av om brukeren oppfatter en kopiering som en bevegelse, og om denne bevegelsen har samme kvaliteter som bevegelsesinntrykket i systemlyden.

- Abstrakt?

Ja, denne systemlyden består av et lydobjekt med typologiske og morfologiske kvaliteter som jobber sammen for å skape inntrykket av dramatikk. Det at lydobjektet har stor masse, er et eksempel på et designvalg som forsterker den abstrakte sannheten.

- Skjematisk?

Nei, egentlig ikke. Det er ikke mulig å tolke noen konkrete verdier ut av lyden, med mindre man ser tegnet 'obstruksjon' som en tallverdi for noe slikt som 'kopieringshastighet er null byte per sekund'. En slik kobling er heller tvilsom.

- Sensorisk?

Ja, lyden kan få oss til å skvette, og attributter i lyden er framhevet for å utnytte dette. Kraftig bass og impulsiv typologi forsterker den sensoriske sannheten. Lydnivået til systemlyden ligger tett opptil det maksimale som kan leveres av brukergrensesnittet, og vil (avhengig av eksterne lydforhold) også forsterke den sensoriske sannheten.

På samme måte som den røde sirkelen med hvit X hadde en betydning som også fantes i skriften, ser vi at lyden står i et lignende forhold. Betydningen 'obstruksjon' formidles gjennom et inntrykk av en bremsebevegelse i lydens naturalistiske sannhetsgrad. Denne betydningen finner vi igjen i skriften 'cannot copy'. Likheten i betydning kan tyde på at vi her ser en forankring, men det er usikkert om denne forankringen har en spesiell retning. På den ene



siden går det an å hevde at vi først reagerer på den sensoriske sannhetsgraden i lyden, deretter leser skriften i feilmeldingen, for så å finne betydningen til systemlyden. På den andre siden går det an å hevde at noe slikt bare inntreffer første gang vi hører denne lydhendelsen. Ved senere inntreff av lydhendelsen gjenkjenner vi lyden og dens betydning, selv om vi skulle ha øynene vendt bort fra dataskjermen i det aktuelle øyeblikket. Når vi så flytter blikket over på feilmeldingen, forankrer vi skriften etter samme regler som i forholdet mellom det grafiske ikonet og skriften.

Jeg ønsker ikke å avgjøre om en av de forankrende betydningsstrømmene nevnt ovenfor er mer riktig enn den andre, men konstaterer at begge kan være gyldige. Samtidig er det viktig ikke å glemme at det også ligger betydning i de grafiske ikonene, og at denne står i et lignende forhold som lydens betydning i forhold til skrift. Dermed er det en kraftig redundans i betydning, som fører til en meningsforsterkning mellom de ulike modalitetene. Dessuten gir en modalitet fordeler som en annen ikke har: lyden kan formidle mening til brukeren selv om brukeren ikke ser på skjermen, og grafiske ikoner kan formidle mening til brukeren selv om lyden skulle være avslått.

I tillegg til at vi har likheter i betydning på tvers av de ulike modalitetene, har vi også forskjeller i betydning. Vi husker at den røde sirkelen med X kunne signalisere fare, men i skriften står det ikke noe om farefulle konsekvenser ved å ha en full harddisk<sup>14</sup>. Lyden på sin side signaliserer dramatikkk gjennom abstrakt og sensorisk sannhetsforsterkning, men i skriften står det ikke noe om faren ved å overse feilmeldingen. Ut fra tolkningen om at avløsning er ulikheter i betydning mellom to modaliteter ([Larsen 1988](#)), slutter jeg meg her til at lydens betydningsstrøm avløser skriften fordi den signaliserer en dramatikkk som ikke kommer fram av skriften i feilmeldingen.

Lydhendelsen Critical Stop formidler også betydning gjennom konnotasjon, og den konnotasjonen jeg ønsker å trekke fram her, er at lyden (og dens kilde, datamaskinen) blir gjort om til et tegn på 'autoritet'. Bak dette resonnementet ligger myten om at 'datamaskiner er dumme', og at de ikke klarer å gjøre annet enn å følge ordre, uansett hvor ulogisk den måtte være. Det er mange tegn i denne feilmeldingen som multimodalt jobber sammen om å skape

---

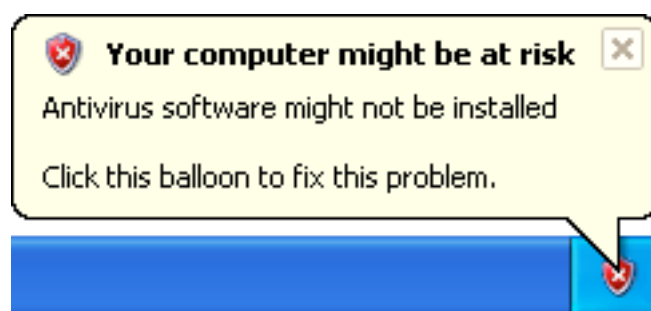
<sup>14</sup> Dagens datamaskiner bør alltid ha noe ledig plass på harddisken, da de baserer seg på virtuell hukommelse og andre midlertidige buffere som fylles opp utenfor brukerens kontroll. Blir harddisken overfylt, er det fare for at datamaskinen slutter å fungere.

en betydning av obstruksjon. Den røde sirkelen med X, den skremmende lyden og den negativt ladede skriften danner alle en multimodal *provenance*-forbindelse mot den gang du skulle på ferie og stappet kofferten full av klær. Det gikk til slutt, på tross av at alle andre familiemedlemmer sa at det ikke gikk. Derfor burde det også gå an å stappe datamaskinen ekstra full med bilder og film, bare du gir på litt ekstra. Men nå viser det seg at datamaskinen ikke gjør annet enn å si ‘stopp’, og plutselig ser du datamaskinen i samme rolle som den personen i innsjekkingsskranken som påpekte at kofferten din var for tung: datamaskinen er blitt en autoritet.

## Hendelse 2 av 4: Notification (Windows XP)

Neste lydhendelse på programmet kan oppstå en kort stund etter at Windows har startet, hvor programmet Windows Security Center ikke ser noen tegn på at antivirusprogramvare er installert og gir brukeren melding om dette. Vi har de samme modalitetene til stede her som i feilmeldingen nevnt ovenfor, så jeg kommer til å gå noe raskere gjennom beskrivelsen av hva vi ser, med bakgrunn i at vi allerede har vært gjennom mange av forholdene.

Nede i høyre hjørne av skjermen dukker det opp en feilmelding formet som en snakkeboble med kremgul bakgrunn, se *Figur 4b*. Ikonet oppe til venstre er bemerkelsesverdig likt den røde sirkelen med hvit X fra forrige lydhendelse, men her er den røde sirkelen erstattet med et rødt skjold. Lukkeruten befinner seg fremdeles oppe til høyre, bare i en annen farge enn tidligere. OK-knappen er ikke til stede, men skriften forteller oss at vi kan klikke på ballongen.



*Figur 4b: Skjermbilde fra Notification (Microsoft 2001)*

Det var en rask gjennomgang av de visuelle symbolene, og nå vil jeg konsentrere meg om lydens sannhetsforsterkning i denne sammenhengen:

- Naturalistisk?

Noe naturalistisk sannhetsgrad, men dette forutsetter en kulturell erfaring med lydeffekter i tegnefilm. Vi ser et ikon som vi tolker som en snakkeboble, og lyden høres ut som en boble som sprekker. Handlingene stemmer ikke overens fordi den ene dukker opp og den andre forsvinner, men temamessig hører de sammen.

- Abstrakt?

Ja. Sammen med det visuelle inntrykket skapes det et tegneserieaktig preg, som importerer en betydning som er mindre seriøs og mer triviell enn hva lydens opprinnelige ‘plutselig bemerkning’-betydning skulle tilsi. Den nye betydningen blir noe nærmere ‘triviell bemerkning’. Lydens interne kvaliteter hermer etter noe trivielt og er således med på å styrke den abstrakte sannhetsgraden.

- Skjematisk?

Nei, det går ikke an å tolke noen målbare data ut fra lydens betydning.

- Sensorisk?

Ja, lyden er i utgangspunktet behagelig på grunn av en lav lydstyrke. Systemlyden blir gjort mer morsom ved at den imiterer en lydeffekt fra tegnefilm. Likevel er virkemidlene ganske nøytrale, og de overdrives ikke for å framprovosere en følelsesmessig reaksjon.

Lydens abstrakte sannhetsforsterkning fører med seg en betydning av trivialitet. Denne betydningen er også til stede i feilmeldingens tegneserieaktige utforming, men den er ikke til stede i det røde grafiske ikonet eller i skriften. Lydens betydning av noe trivielt *forankrer* tilfellene av ordet ‘might’ i skriften, og den *avløser* det røde ikonet som vi da ikke lenger tolker som den obstruksjonen fra forrige lydhendelse, men heller et symbol uten direkte forbindelse til feilmeldingens innhold. Det røde skjoldet med hvit X er egentlig et symbol som hører til Windows Security Center, men det forstår vi ikke før vi eventuelt klikker på snakkeboblen og går videre. Det er ikke noe ved lydens karakter som tyder på at vi kan klikke på snakkeboblen (husk: boblen er ifølge lyden allerede sprukket), så betydningen ‘klikkbarhet’ blir bare formidlet gjennom skriften. Imidlertid kan det hende at brukeren allerede har avfeid lydhendelsen som noe trivielt og derfor ikke leser skriften. Etter noen sekunder forsvinner også snakkeboblen og feilmeldingen som en helhet.

Vi har også her et konnoterende element, som jeg ser som tegnet på ‘mas’ (at datamaskinen maser på brukeren). En fare ved å bruke morsomme virkemidler slik som her hvor det brukes

virkemidler fra tegneserieverdenen, er at det morsomme kan bikke over til å bli påtrengende og plagsomt (Sharp et al. 2007:184). Noen lesere husker kanskje Office-assistenten, den levende binders i Microsoft Office som med glimt i øyet fant det for godt å påpeke at det så ut som om brukeren prøvde å skrive en liste, og kanskje kunne trenge hjelp til dette. De gangene dette normalt søte og morsomme virkemiddelet inntraff for ofte eller i uønskede situasjoner, ble det plagsomt. Forklart ut fra sannhetsforsterkning går det vel an å si at graden av sensorisk sannhet vil kunne være like sterk uansett hvilke følelser som denne forsterkningen får fram. Hvis myten er at morsomme kommentarer ikke får deg til å jobbe bedre eller raskere, blir lyden og dens andre medfølgende modaliteter til sammen et tegn på ‘mas’. Dette tegnet forsterkes ved repetisjon, og hvis du velger å se bort fra feilmeldingen, vil den dukke opp på ny ved neste oppstart av datamaskinen. Dermed gjentas meningsformidlingen helt til vi overser lydens trivielle betydning og tolker skriften slik at vi kan klikke oss videre.

### Hendelse 3 av 4: Device Connect (Windows XP)

Denne lydhendelsen skjer en stund etter at vi har gjort handlingen i brukerverdenen. I dette tilfellet har jeg skrudd på en USB-harddisk som var koblet til datamaskinen, og etter noen sekunder hører jeg lyden til Device Connect og ser en feilmelding som i Figur 4c. Det visuelle har store likheter med den forrige hendelsen jeg analyserte, vi har den samme gule snakkeboblen med lukkerute, ikoner og skrift. Det røde skiltet med hvit X er her byttet ut med en hvit snakkeboble med påskrevet i-bokstav, et symbol som vi kjenner inn fra andre situasjoner i hverdagen hvor det betyr ‘informasjon’.



Figur 4c: Skjerm bilde fra Device Connect (Microsoft 2001)

- Naturalistisk?

Nei, ikke hvis man sammenligner med handlingen i brukerverdenen. Ikke bare er det få eller ingen tilkoblinger av utstyr som låter slik i naturen, men det er også en vesentlig tidsforsinkelse mellom når vi utfører handlingen og når vi hører og ser meldingen. Handlingen i datamaskinens virtuelle verden er dessuten skjult og ukjent, og objektene den behandler er

også skjulte og ukjente. Da har den virtuelle verdenen egentlig ingen egenskaper som lyden kan knyttes til. Naturalistisk sannhetsforsterkning forekommer derfor ikke.

- Abstrakt?

Ja. Intervallspranget kompenserer for de manglende naturalistiske forbindelsene. Graden av abstrakt sannhetsforsterkning er mellomsterk. Vi snakker ikke her om en lyd som prøver å formidle noe om livets store spørsmål, men den abstrakte sannheten forsterkes multimodalt ved hjelp av symmetri mellom forskjellige bevegelsesinntrykk, og lyden spiller på bevegelsesinntrykk gjennom metaforer.

- Skjematisk?

Bare såvidt. Spranget fra et lavt til et høyere nivå kan bety 'vellykket', men vi vet ikke noe om hvor vellykket det er. Maskinen fant tydeligvis en ny USB-disk, men hva heter den og hvor stor kapasitet har den? Dette er egenskaper som kan høre til lydhendelsen, men som lyden her ikke formidler.

- Sensorisk?

Ja, "glad er opp"-metaforen spiller på følelser. Likevel er det liten grad av forsterkning – det er ikke akkurat barnelatter og sirkusmusikk vi hører her. Virkemidlene er til stede, men de forsterkes ikke.

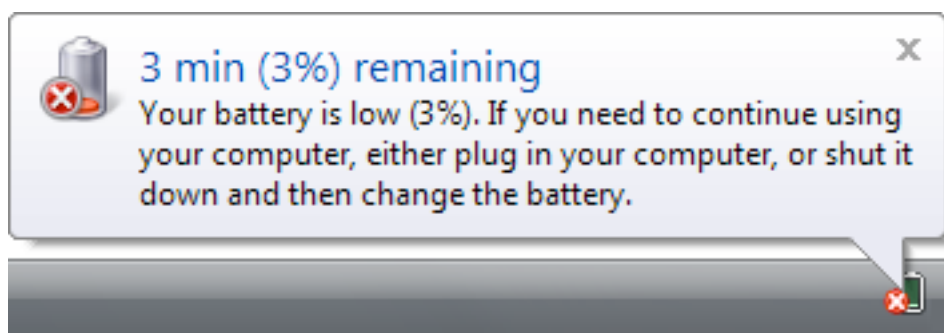
I motsetning til Notification-hendelsen som vi nettopp har vært gjennom, er det her store likheter i betydning på tvers av alle modaliteter. Lyden bruker sensorisk forsterkning til å skape en positiv følelse hos brukeren, feilmeldingen har en mild gulfarge, ikonet oppe til venstre i feilmeldingen er et ikke rødt, skarpt og skjærende slik som det var i Notification, og skriften inneholder ingen advarsler eller andre negative betydninger. Graden av sensorisk forsterkning i lyden er derimot liten, så hele lydhendelsen forblir nøytral og mild.

Mangelen på naturalistisk sannhetsgrad hemmer ikke lydens meningsformidling, da den abstrakte sannhetsgraden stemmer såpass godt overens med betydningen fra de andre modalitetene. Et viktig poeng er likevel at den abstrakte sannhetsgraden er avhengig av å få betydning fra andre modaliteter. Hvis skriften ikke hadde forankret den abstrakte betydningen i lyden hvor toneintervallet betyr 'tilkobling', ville det vært vanskelig å vite hvilken brukerhandling som lyden skulle knyttes til. Dette er spesielt aktuelt i denne situasjonen hvor det tok noen sekunder fra jeg skrudde på harddisken og til systemet gjenkjente den som en ny enhet. Enda viktigere blir det i de situasjonene hvor denne lydhendelsen ikke er brukerskapt,

men snarere er et resultat av endringer internt i datamaskinen, for eksempel en driveroppdatering. Lyden inneholder ingen betydning om hva slags enhet det dreier seg om. Informasjonen kommer enten skriftlig eller forstås som et hendelsesforløp.

## Hendelse 4 av 4: Battery Critical (Windows Vista)

Vi avslutter denne multimodale analysen ved å gå over til et eksempel fra Windows Vista, som inntreffer når batterinivået er svært lavt, på grensen til at maskinen dør. Resultatene av denne hendelsen vil variere mellom ulike datamodeller. Noen datamaskiner forsetter som før helt til de dør, mens andre går over i dypdvale<sup>15</sup> noen sekunder etter at du har hørt Battery Critical og sett en feilmelding som i *Figur 4d*. Feilmeldingen ligner snakkeboblen som vi kjenner fra Windows XP, den har litt andre farger, men plasseringen av de ulike elementene er som før.



*Figur 4d: Skjerm bilde fra Battery Critical (Microsoft 2006)*

- Naturalistisk?

Nei, det er ikke noe i lyden som gjenspeiler lyden av lavt batteri. Et eksempel på en naturalistisk framstilling kunne vært musikk som ble spilt saktere og saktere, og ble dypere og dypere i frekvens parallellt med at den gikk saktere.

- Abstrakt?

Ja, og det skyldes to ting. For det første er det lite som stemmer mellom lydens interne betydning og handlingen at maskinen går tom for batteri, som diskutert i den deskriptive analysen. For det andre går det noe tid fra vi hører lyden til vi merker konsekvensen: at datamaskinen går i dypdvale eller slår seg av. Lyden blir derfor et symbol på noe som skal skje i nær, men ukjent framtid, men handling-lyd-slektskapet gjelder ikke fordi lyden kommer før handlingen som den representerer.

---

<sup>15</sup> Dypdvale betyr at datamaskinen skriver minnet sitt til harddisk, for deretter å slå seg av. På den måten kan man fortsette å jobbe videre så snart maskinen får strøm igjen.

- Skjematisk?

Nei, men dette er en lyd som absolutt kunne vært mer virkningsfull med en skjematisk sannhet. Det hadde vært praktisk å vite hvor mye av batteriet som var igjen bare ved å høre på en attributt i selve lyden.

- Sensorisk?

Nei. Det er lett å tro at vi her er vitne til en sensorisk sannhetsforsterkning, siden lyden både begynner og slutter mykt og spilles av på et lavt og behagelig lydnivå. Imidlertid er ikke dette attributter som overdrives, og overdrivelse kreves for at det skal foregå en sensorisk

sannhetsforsterkning – “amplified beyond the point of naturalism” (Van Leeuwen 1999:161).

Lyden for Battery Critical er rett og slett ganske statisk og fri for interne kontraster, og har derfor liten sensorisk sannhetsforsterkning.

Til tross for at vi har lignende visuell symbolbruk som i Critical Stop og at skriftens betydning kommer i form av et beordrende språk, reduseres disse negative virkemidlene av lydens manglende sannhetsforsterkning. Lydens betydning kommer her i konflikt med feilmeldingens andre virkemidler. Fordi lyden er myk og svak, vies ikke feilmeldingen så mye oppmerksomhet som man skulle tro de visuelle virkemidlene la opp til. I den grad det finnes multimodale forbindelser til lyden, eksisterer disse som en indeksikalsk (ikke-naturlig) forbindelse hvor de visuelle inntrykkenes betydning forankrer den abstrakte sannhetsgraden i lyden. Derfor spiller lyden en viktig rolle i Battery Critical, men en noe uvanlig rolle fordi den først og fremst bidrar til å redusere brukerens oppmerksomhet, og på den måten svekkes også meningsformidlingen.

## 4.3 Problemstilling 2: Systemlydenes betydningsstrøm

| Hendelse                 | Naturalistisk                         | Abstrakt                              | Skjematisk          | Sensorisk                 |
|--------------------------|---------------------------------------|---------------------------------------|---------------------|---------------------------|
| Critical Stop (XP)       | Middels, bare som bevegelsesinntrykk  | Middels, fra lydens interne betydning | Lav, bare i teorien | Høy, bevisst overdrivelse |
| Notification (XP)        | Middels, kulturelt avhengig           | Middels, fra lydens interne betydning | Ingen               | Middels, overdrives ikke  |
| Device Connect (XP)      | Lav, bare i teorien. Tidsforsinkelse. | Høy, forsterkes multimodalt           | Lav, bare i teorien | Lav, men med potensiale   |
| Battery Critical (Vista) | Lav, vanskelig å argumentere for      | Lav, indeksikalsk forbindelse         | Ingen               | Lav, veldig nøytral       |

Tabell 4a: Sannhetsgrader i analyseobjektene

Den mest åpenbare tendensen som kommer ut av å sammenligne sannhetsgradene i de ulike lydhendelsene, er at det praktisk talt ikke eksisterer noen skjematisk sannhetsgrad i lydene. Dette fraværet skyldes ikke mangel på egenskaper som kan beskrives skjematisk: Critical Stop kunne inneholdt informasjon om hva slags objekt som stoppet, Device Connect kunne enten beskrevet den tilkoblede enheten eller en av dens egenskaper, og Battery Critical kunne gitt brukeren informasjon om hvor mye strøm som var igjen i batteriet. Jeg ser umiddelbart en grunn til dette, og den er av teknisk karakter. Systemlydene skal ikke bare skal brukes opp mot den konkrete feilmeldingen som jeg har analysert, men de skal kunne gjelde for en hel rekke situasjoner. Notification-lyden må for eksempel gjelde for alt fra reklame til sikkerhetsadvarsler. Systemlydene er ferdiginnspilte lydfiler som er designet på forhånd og har derfor ingen mulighet til å innrette seg etter variabler i lydhendelsen.

Den andre tendensen jeg ser, er at det virker som om sannhetsforsterkning kan smitte over fra en sannhetsgrad til en annen. Dette kan skyldes at kategoriene ikke er godt nok definert, men det kan også skyldes at betydning beveger seg på tvers av sannhetsgradene. For eksempel har de to lydene med høyest sensorisk sannhetsgrad også en middels høy naturalistisk og abstrakt sannhetsgrad. Kanskje de sansene som stimuleres hos lytteren gjennom sensorisk sannhetsforsterkning vekker en iboende egenskap i oss til å finne betydning i lydene? Unntaket er Device Connect, som har høy abstrakt sannhetsgrad til tross for at den sensoriske sannhetsgraden er lav. Her blir imidlertid det abstrakte importert multimodalt, så denne verdien kan holdes utenfor statistikken for tendenser på tvers av sannhetsgradene. Kanskje den lyden til og med er unntaket som bekrefter regelen.



## Forbindelser med andre modaliteter

| Hendelse                    | Forankring   | Avløsning   | Konnotasjon      |
|-----------------------------|--|---|------------------|
| Critical Stop<br>(XP)       | Fra lyd til skrift:<br>handling (stopp)                  | Fra lyd til skrift:<br>oppmerksomhet (dramatisk)        | Autoritet        |
| Notification<br>(XP)        | Fra lyd til skrift og bilde:<br>oppmerksomhet (trivielt) | - ingen funnet -  | Mas              |
| Device Connect<br>(XP)      | Fra skrift til lyd:<br>handling (tilkobling)             | - ingen funnet -  | - ingen funnet - |
| Battery Critical<br>(Vista) | Fra skrift til lyd:<br>handling (batteritap)             | Fra lyd til skrift og bilde:<br>oppmerksomhet (uviktig) | - ingen funnet - |

Tabell 4b: Forbindelser mellom lyd og andre modaliteter i analyseobjektene

Som jeg fant ut i den multimodale analysen av Critical Stop, er måten vi leser feilmeldingen på avhengig av vår erfaring med å lese feilmeldinger. Derfor er de verdiene jeg har ført inn i **Tabell 4b** bare noen få eksempler på forbindelser mellom modalitetene. Jeg klarte riktignok ikke å finne spor av avløsning eller konnotasjon for alle lydhendelsene, men under forankring har jeg måttet redusere funnene mine til enkelte poeng som jeg synes var viktig å få fram.

De to første lydhendelsene hadde en merkbar naturlig sannhetsgrad, og de preges av at lyden forankrer betydning i de andre modalitetene. Når vi hører en systemlyd, får vi gjennom lydens handling-lyd-slektskap inntrykket av at lyden ble skapt av en handling, og vi får en forventning om at noe i brukergrensesnittet har skapt lyden. For å finne ut hva som ligger bak denne handlingen, retter vi blikket mot betydning i feilmeldingen. Er det likheter til stede, forankrer lyden de modalitetene med lignende betydning. I Critical Stop lå likheten i bremsebevegelsen, det grafiske ikonet og skriften 'cannot copy'. I Notification lå likheten i lyden av en boble som sprakk og at vi så en snakkeboble som dukket opp. Disse var to forskjellige handlinger, men de var relaterte i tema og dukket opp samtidig. Derfor virker koblingen naturlig. I Device Connect lå det derimot en vesentlig forsinkelse mellom brukerens handling og brukergrensesnittets lydhendelse. Når det tar lang tid mellom vi når gjør en handling og når vi hører at den blir gjort, er ikke handling-lyd-slektskapet spesielt sterkt. Vi blir derfor oppmerksomme på feilmeldingen, leser dens visuelle betydning, og overfører den betydningen til den handlingen vi gjorde tidligere. Når det er høy naturlig sannhetsgrad, får vi mye betydning fra lyden, mens høy abstrakt sannhetsgrad gjør at vi importerer betydning i lyden fra de andre modalitetene.

Når det gjelder avløsning og konnotasjon, virker det som disse meningsutvekslingene ikke alltid er lette å oppdage. Mangel på konnotasjon kan godt forklares ut fra at den er avhengig av en spesiell kulturell bagasje, mens mangel på avløsning kan skyldes at det semiotiske språket til lydikonene ikke er tydelig nok til å kunne danne en betydning som peker seg ut i forhold til andre samtidige inntrykk. Et interessant moment er at de systemlydene hvor jeg påviste konnotasjon, også er de systemlydene med høyest sannhetsforsterkning fordelt over flere sannhetsgrader. Dette kan være et hint på at sannhetsforsterkning aktiverer multimodale forbindelser i oss, som gjør det lettere å omsette betydning og behandle den på nytt. Jeg finner imidlertid ikke nok data i analyseobjektene mine til å underbygge denne påstanden.

## **Oppsummering**

Systemlydenes betydningsstrøm er svært sammensatt, og det er bare noen få poeng jeg har klart å trekke fram som tendenser. Ingen av analyseobjektene benyttet seg av skjematisk sannhetsforsterkning, fordi tekniske forhold rundt avspilling og lyddesign ikke tillater det. En høy naturlig sannhetsgrad kan føre til at betydning overføres fra lyden til skrift og bilde, mens en høy abstrakt sannhetsgrad kan føre til at betydning overføres fra skrift og bilde til lyd. Det virker også som om sannhetsforsterkning kan smitte over fra en sannhetsgrad til en annen, først og fremst i lyder hvor den sensoriske sannhetsgraden er høy.

Mange av systemlydene har et iboende potensiale til å forsterke en eller flere av sannhetsgradene, men det var bare i noen få tilfeller at dette potensialet ble utnyttet. Det virker som om det kan ligge en frykt mot å forsterke betydningen i lydene, særlig i lyden til Windows Vista. I neste kapittel skal jeg se problemstillingene på avstand og undersøke mulige årsaker til hvorfor det er slik.

## 5. Kontekstualisering

---

**D**e to analysekapittelene mine har gitt meg mye å snakke om, inkludert at det kan stilles spørsmål om årsaker til at systemlydene er blitt designet slik som de er. For best å kunne besvare et slikt hvorfor-spørsmål, skal jeg i dette kapittelet se analyseresultatene mine i forhold til konteksten utenfor selve brukergrensesnittet. Først vil jeg gjøre rede for et syn på systemlyder som kommer fram hos Microsoft, deretter se på et motsvarende forskersyn og sammenligne disse ulike tilnærmingene. Slik tenker jeg å komme fram til hvor utviklingen av systemlyder står, og hvordan dette påvirker valgene som gjøres når systemlyder designes for et brukergrensesnitt.

### 5.1 Microsofts systemlydformål

Microsoft har på sine nettsider for utviklere laget en samling med retningslinjer som skal veilede utviklere av Windowsprogrammer til å skape en brukeropplevelse som stemmer overens med Windows Vista: “The Windows Vista User Experience Guide” ([Microsoft 2008](#)). Det vil være naturlig å tro at om ikke de samme, så i hvert fall lignende retningslinjer har eksistert internt hos Microsoft under utviklingen av deres eget operativsystem, og at dette har vært noe som Microsofts lyddesignere også har tatt hensyn til da de lagde systemlydene til Windows Vista. En grei kontekstualisering kan derfor være å se hvordan Microsofts retningslinjer stemmer overens med analyseresultatene mine. [Microsoft \(2008:704\)](#) ser fire formål ved å bruke systemlyder:

- Opplysning etter at datamaskinen har utført en handling, som lyden av nyankommet e-post
- Tilbakemelding etter at brukeren har utført en handling, som å spille av lyd etter at brukeren har justert lydstyrken til operativsystemet
- Branding, enten ved å lage et eget lyduttrykk som forsterker oppfattelsen av et dataprogram som et helhetlig produkt, eller ved at dataprogrammets lyduttrykk gir en opplevelse av å være innlemmet i medieproduktet Windows Vista
- Underholdning, særlig i spill eller i opplæringsprogramvare

*Opplysning* var hovedformålet til alle lydhendelsene som jeg tok for meg i analysen. Alt fra beskjedne Notification-hendelser til alarmerende Critical Stop inntraff etter en handling som hadde hendt i datamaskinens virtuelle verden. Et gråsonceeksempel finner vi i Device Connect, som både kunne inntreffe på grunn av handlinger i datamaskinen (oppdaterte drivere) og på grunn av handlinger fra brukeren (tilkobling). Til en viss grad kan det argumenteres for at flesteparten av datamaskinens handlinger oppstår i brukerverdenen som følge av en eller annen brukerhandling, men jeg ser ikke hensikten i å finne ut av det for å avgjøre om Microsoft har holdt seg til sine retningslinjer.

Selv om systemlydene som jeg analyserte ga både *opplysning* og *tilbakemelding* til brukeren, var det svært sparsomt med informasjon informasjonen om *hva* som hadde hendt. I den multimodale analysen fant jeg ut at mesteparten av den detaljerte informasjonen om ‘hva, hvordan og hvorfor’ ble gitt skriftlig, og at lyder ble spilt av for å formidle betydning som var underlagt dette (til tross for noen hederlige unntak). Jeg nevnte ovenfor at Microsofts eksempel på ‘tilbakemelding’ kan være å spille av lyd etter at brukeren har stilt inn operativsystemets lydstyrke. Til en slik handling bruker Windows lydhendelsen Default Beep. Når lydstyrken justeres, spilles det av en lyd når brukeren har justert lyden til et visst nivå. Avspillingsvolumet til lyden gir da en indikasjon på hvor kraftig lydnivå brukeren har stilt inn, og dette gjøres gjennom en skjematisk sannhetsforsterkning (Van Leeuwen 1999) hvor lydnivået er en egenskap ved lyden som brukeren kan omsette til konkrete verdier. Bortsett fra dette eksemplet, finner jeg imidlertid få eller ingen andre eksempler på bruk av skjematisk sannhetsforsterkning i analyseobjektene mine. Det kan virke som om Microsoft skiller mellom at ‘opplysning’ skal tilkalle oppmerksomhet og at ‘tilbakemelding’ skal inneholde informasjon, og å forstå inndelingen som et slikt skille er etter min mening uheldig. Skal systemlydene inneholde informasjon som brukeren kan tolke, bør dette kunne gjelde uansett lydens formål.

*Branding* er et stort felt som jeg til nå har holdt utenfor oppgaven, og jeg skal heller ikke gå for mye inn på det her. Kort forklart er branding en markedsføring av ens produkter ved bevisst repetering av enkelte symboler, eller “emotional positioning of a product as perceived by its customers [...], the attributes of the product experience itself” (Microsoft 2008:712). Windows Vistas oppstartslid er et eksempel på bevisst bruk av branding (ibid.:704), og det er en stor industri bak det å lage kjenningsmelodier, deriblant oppstartslider. Microsoft benyttet seg av

den britiske artisten Robert Fripp til å lage oppstartsslyden til Windows Vista, og de har tidligere brukt artister som Brian Eno og Rolling Stones til samme arbeid ([Torre 2006](#)).

Som nevnt i [Kapittel 3.1.4](#) har jeg ikke analysert oppstartsslyden til Windows. En ting jeg likevel kan påpeke, er at de kvalitetene som Microsoft ønsket i oppstartsslyden ([Torre 2006](#)) stemmer overens med de kvalitetene som Microsoft ønsker at Vista-lydene skal ha ([Microsoft 2008:705](#)). Det er derfor naturlig å tro at et av formålene ved lyddesignet til systemlydene i Windows Vista er å forlenge branding som skapes av oppstartsslyden ved å ha lyder som på repeterende vis påminner brukeren at 'dette er medieproduktet Windows Vista'. Ryktene før Vista ble sluppet ut om at Microsoft ikke ville legge til rette for at brukerne kunne skru av oppstartsslyden ([Reimer 2006](#)) styrker teorien om at Microsoft tar branding-momentet seriøst.

*Underholdning* er et formål som utnytter lydens affektive egenskaper, med andre ord lydens evne til å skape (fortrinnsvis gode) følelser hos lytteren. Her går det an å trekke en parallell til den sensoriske sannhetsforsterkningen nevnt i [Kapittel 4](#), som var et virkemiddel i lyd som forsterket følelsene hos lytteren ved å overdrive attributter i lyden. Vi husker for eksempel lyden Notification i Windows XP som brukte virkemidler fra tegneserieverdenen til å skape en litt underholdende effekt. Microsoft har lang erfaring fra spillbransjen, og i de siste årene har de også laget egne spillkonsoller. Det ser imidlertid ut til at få av systemlydene til Microsoft Windows er laget nettopp med underholdning i tankene, og at det heller er opplysning og branding som har fått førsteprioritet.

Ved underholdningsformålet er det verdt å stille spørsmål ved hvor grensen går mellom underholdning og en god brukeropplevelse. Er underholdning nødt til å være spennende, og er en god brukeropplevelse nødt til å være kjedelig? Betyr underholdning at man er nødt til å legge merke til lydene bevisst, mens en god brukeropplevelse betyr at lydene tolkes underbevisst underbevisst? God og brukervennlig design kan påvirke følelsene våre, og design som tar hensyn til følelsene våre skaper tilsvarende et inntrykk av god og brukervennlig design ([Norman 2004](#)). Det er ikke sikkert at det finnes noen spesielle grenser som kan brukes til å kategorisere lyder som har underholdning som formål, men at underholdning i stedet er en egenskap som kan brukes positivt innenfor flere kategorier.

## 5.2 Microsofts faremomenter

Etter å ha omtalt bruk av systemlyder som noe overordnet positivt, endrer Microsoft tonen videre i retningslinjene sine, og kommer fram til at systemlyder må brukes varsomt:

“Despite the benefits of sound, appropriate use of sound requires significant restraint – to do otherwise can make a program annoying and distracting. Users will turn off their sound completely if they become annoyed by frequent, repetitive, jarring, disrupting, poorly designed sounds; in part this is because by its very nature, sound demands attention and is hard to ignore. [...] Because the drawbacks of using sound can easily outweigh the benefits, use sound only when there is a clear advantage. When in doubt, don’t use sound.” (Microsoft 2008:704)

Sitatet ovenfor preger store deler av hva Microsoft har å si i om lyd i sin veiledning. Nedenfor kommer jeg til å undersøke hvilke faktorer som kan ligge bak at Microsoft konkluderer slik de gjør. Jeg ser tre områder hvor uromomenter kan komme inn og gjøre bruk av systemlyder problematisk: Tekniske og miljømessige hensyn, tilrettelegging for funksjonshemmede og til slutt smak og kultur.

### 5.2.1 Tekniske og miljømessige hensyn

Microsoft nevner fire situasjoner hvor lyd kan være ugunstig, og alle ligger under det jeg vil kalle tekniske og miljømessige hensyn: Systemlyder i bråkete miljøer, systemlyder i stille miljøer, datamaskiner uten høyttalere og profesjonelle servere (Microsoft 2008:704).

I et bråkete miljø vil systemlyder kunne drukne i resten av lydbildet, og derfor ikke bli hørt. Det samme gjelder datamaskiner som ikke har høyttalere eller annet avspillingsutstyr. Hvis systemlydene i slike situasjoner inneholder betydning som både er *kritisk* for brukerens forståelse og at denne betydningen samtidig *ikke kommer fram* gjennom andre modaliteter, vil brukeren dermed kunne gå glipp av av nødvendig informasjon. Microsoft baserer seg på denne oppfatningen og legger den til grunn for sitt eget lyddesign, ved at systemlydene deres ikke formidler mange spesifikke detaljer om lydhendelsen. Dette vises i analyseresultatene mine, hvor den mest spesifikke informasjonen så ut til å formidles gjennom skrift.

Stille miljøer omfatter steder som bibliotek, lesesaler, styremøter og alle andre situasjoner hvor uønsket lyd lett blir pinlig. Her tas det hensyn til at lyder ikke skal forstyrre menneskers konsentrasjonsevne. I vår stressende hverdag er dette naturligvis et viktig hensyn å ta, men spørsmålet er hvordan. Skal brukeren selv være ansvarlig for å dempe lyden på datamaskinen, eller bør hensyn tas av programvareutvikleren slik at lyd bare brukes der det er absolutt nødvendig? Hvis en skal følge Microsofts råd og gjøre det sistnevnte, kan det bli lettere for situasjoner å oppstå hvor en bruker glemmer å skru ned lydstyrken eller tror at lyden allerede er dempet, bare for å få seg en overraskelse når brukergrensesnittet plutselig gir lyd fra seg. Et annet moment i denne diskusjonen er i hvor stor grad ulike lyder krever vår oppmerksomhet. Lyd oppfattes både via bevisste ‘ovenfra og ned’-prosesser og via mer underbevisste ‘nedenfra og opp’-prosesser i hjernen (Snyder 2000, Branigan ifølge Maasø 1995:28-29). Hvis lyder designes slik at betydning kan formidles mer via ‘nedenfra og opp’-prosessering og mindre via virkemidler som krever mye oppmerksomhet (for eksempel verbal kommunikasjon), kan det bli plass til flere lyder i lydbildet, i motsetning til dagens avspillingsløsning som bare spiller en systemlyd om gangen.

Profesjonelle servere er datamaskiner som kan stå i et helt annet rom enn brukeren, og som utfører handlinger automatisk, opp mot hundrevis av oppgaver på en gang (Microsoft 2008:704). Derfor er det ikke alltid det passer å bruke systemlyder til å beskrive handlinger i en servers virtuelle verden, fordi de utføres automatisk uten direkte tilsyn og med så høy frekvens at individuelle lydhendelser og deres betydning kan bli grøtete. Denne situasjonen regner jeg som et spesialtilfelle av ‘stille miljøer’ nevnt ovenfor. Det er klart at jeg som bruker ikke er interessert i å høre en stemme si ‘Operation completed!’ på samme måte hvert femte sekund. På den annen side kan jeg være veldig interessert i å vite at linjekapasiteten min er oppbrukt fordi en spesiell oppgave tar alle ressursene. En lydmessig måte å representere noe slikt på kunne for eksempel være en tone full av lydobjekter med iterativ typologi hvor hver iterasjon gjenspeiler ledig linjekapasitet, og hvor morfologien peker på hva slags informasjon som strømmer gjennom. Derfor mener jeg at profesjonelle servere ikke trenger å være miljøer hvor utviklere ser bort fra å bruke lyd, det er mer et spørsmål om hvordan lyd implementeres i brukergrensesnittet.

Et annet forhold som jeg vil inkludere innunder tekniske og miljømessige hensyn, er hvis systemlyder brukes på langsomme datamaskiner eller utstyr som ikke er optimert for

lydavspilling. I dag har personlige datamaskiner mer enn nok regnekraft til å behandle lyd slik den er implementert i Windows-brukergrensesnittet, så dette problemet var større før. Likevel finner vi eksempler på små elektroniske enheter, for eksempel mobiltelefoner som bruker Windows Mobile som operativsystem, som kanskje ikke inneholder avanserte lydkort grunnet ønsket om å spare pris, plass og strømforbruk. Avspilling av for mye lyd i et slikt operativsystem vil kunne gå ut over andre ressurser som regnekapasitet, responstid og strømforbruk. Dette er forhold som gjelder uavhengig av hvor meningsbærende lyden er. En mulig måte å bøte på dette problemet er ved å se seg selv tilbake i tid. Den gang brukte spillkonsoller en type lydframstilling som ikke var like ressurskrevende som dagens metoder, og som likevel kunne bli sett på som kreativ og meningsbærende.

### **5.2.2 Tilrettelegging for funksjonshemmede**

Databrukere med ulike funksjonshemminger er en viktig brukergruppe å tenke på når et brukergrensesnitt skal utformes. Det er et poeng at de funksjonshemmede ikke går glipp av informasjon som er kritisk for å kunne forstå brukergrensesnittet, og at de heller bør få tilgang til informasjon gjennom flere midler enn det en funksjonsfrisk bruker vil kunne oppfatte som tilstrekkelig.

Microsofts løsning på å legge lyd til rette for funksjonshemmede, i dette tilfellet de hørselshemmede, er den samme løsningen som Microsoft foreslo for de miljømessige hensynene som jeg nevnte ovenfor: unngå å legge betydning i lyd som ikke kommer fram gjennom andre modaliteter ([Microsoft 2008:704](#)). Dette designvalget begrenser mengden betydning som legges i lyden, og er i strid med ideen om at funksjonshemmede bør få informasjon gjennom flere modaliteter for å minske deres avhengighet av en sans alene. Det er naturlig å tro at en synshemmet bruker kan ha vel så mye utbytte av lyder med klar betydning som en funksjonsfrisk bruker, men dagens løsning for synshemmede i Windows er at informasjon leses opp verbalt fra skjermen. Hensynet overfor hørselshemmede svekker dermed muligheten til å integrere systemlyder kreativt som meningsbærere for synshemmede, en konsekvens som jeg mener fungerer delvis mot sin hensikt.



### 5.2.3 Smak og kultur

“Sound can be a polarizing design element. For every user who loves a sound, there will be many who dislike it.” (Microsoft 2008:705)

“Don’t choose sounds that require localization. You can achieve this by using sounds that don’t use speech or have culturally-dependent meanings or connotations.” (ibid.:709)

Denne oppgavens innhold understreker det paradoksale i det nederste av sitatene ovenfor, da jeg i **Kapittel 2.3.2** forklarte hvordan et budskap alltid vil konnotere avhengig av brukerens kultur. Det øverste sitatet styrkes imidlertid av et slikt syn. Jeg skal ikke gå inn på en diskusjon om følelser og smak, men bare legge fram et moment som er viktig å tenke på i utviklingsprosessen:

Det er ikke mulig å blidgjøre alle, selv om det er lov å prøve. Dårlige erfaringer fra tilbakemeldingen til enkelte interessegrupper kan skape en frykt mot å tørre å satse på nye og radikale lydløsninger, og det kan også føre til at noen ideer som nesten fungerte, blir fjernet fra den endelige utgaven av et program eller operativsystem. Det er ikke engang nødvendig å få negativ tilbakemelding for at en slik fjerning skal skje, da det kan eksistere frykt blant utviklere eller deres overordnede som går ut på at man er redd for at ens produkter skal bli upopulære hos noen hvis man lager ubehagelige lyder. Hvis et utviklerteam spiller en lyd foran hundre personer og ønsker å finne ut om lyden er behagelig eller ikke, vil antagelig noen synes at den er ubehagelig. Er det nødvendig å endre lyddesignet av den grunn?

Poenget mitt er å understreke at en lyd kan alltid være behagelig eller ubehagelig, men det viktige er om den er *passende*. Hvis den er upassende, trenger ikke dette skyldes lyden, men det kan ligge en rekke andre forhold i brukergrensesnittet som påvirker vår forståelse av lyden. Hvis en bruker gir tilbakemelding om at en lyd er ubehagelig, er det viktig å huske på at det vel så godt kan være situasjonen som lyden oppstår i, som skaper ubehag hos brukeren. Det vil være lettere å avgjøre disse om forholdene skyldes lyden i seg selv, hvis de som tar avgjørelsene om å beholde lydene kjenner til forskning som omtaler menneskets lydforståelse.

## 5.3 Et forskersyn

Ved å sammenligne funnene mine med Microsofts syn på systemlyder, har jeg sett analyseobjektene mine i lys av de motivene som lå bak deres utforming. En annen måte å kontekstualisere funnene mine på, er å sammenligne dem med noen syn som forskere har på systemlyder i brukergrensesnitt. Forskningen som jeg drar fram her, er over 20 år gammel, men dens synspunkter farger målene til moderne forskning i dag, inkludert teori som jeg har lagt til grunn for denne oppgaven ([Rocchesso og Fontana 2003](#)).

### 5.3.1 Visjoner om lyd i brukergrensesnitt

Hvis man skal begrunne formålet ved å ha systemlyder i et brukergrensesnitt, kan det være greit å begynne med Donald Normans tilbakemeldingsprinsipp, “The Principle of Feedback” ([Norman \[1988\] 2002:27](#)). Her nevnes det paradokset at etter hvert som teknologien blir mer abstrakt, får den økt funksjonalitet samtidig som brukeren får mindre tilbakemelding på hva som faktisk skjer. Norman eksemplifiserer dette ved å nevne hvordan graden av lydmessig tilbakemelding har blitt redusert i overgangen fra analoge til digitale telefonapparater ([loc.cit](#)). Dette er til en viss grad en riktig påstand, fordi hendelsene som foregår i et brukergrensesnitt blir mer og mer abstrakte og dermed vanskeligere å knytte direkte til noe lydskapende i brukerverdenen. Som en konsekvens kan det være aktuelt å ønske seg mer lydmessig tilbakemelding i brukergrensesnittet til datamaskiner, men spørsmålet er hvordan dette kan gjøres på en måte vi oppfatter som naturlig, uten at lydene krever for mye oppmerksomhet.

#### Direkte manipulering

Det er lett å se for seg et ideelt brukergrensesnitt som en situasjon hvor vi slipper å måtte tenke på brukergrensesnittet i seg selv, og heller fokusere på det som brukergrensesnittet formidler. Donald Norman beskriver dette gjennom en prosess han kaller direkte manipulering:

“When I use a direct manipulation system [...] I do think of myself not as using a computer but as *doing the particular task*. The computer is, in effect, invisible. The point cannot be overstressed: make the computer system invisible.” ([Norman \[1988\] 2002:185](#), min utheving)

Et godt brukergrensesnitt som benytter direkte manipulering, er hvor brukeren ikke er oppmerksom på brukergrensesnittet, men heller forestiller seg at objekter i datamaskinens virtuelle verden manipuleres direkte. I stedet for at vi ber *datamaskinen* flytte på tegn fra et sted til et annet, er det *vi* som klipper ut og limer inn tekst. Som Norman selv nevner, er det ikke alle handlinger hvor det passer å benytte direkte manipulering, for eksempel handlinger som krever mye automatikk (Norman [1988] 2002:184). Fremdeles, ønsket er der om at brukergrensesnittet skal være så gjennomiktig og usynlig som overhodet mulig. Dette er en idé som lett smitter over til andre forskere, også innenfor forskning på systemlyder.

## Lydikoner

William Gaver baserer seg på Normans visjoner om gjennomiktige brukergrensesnitt (Gaver 1989), og kombinerer dem med økologisk persepsjonsteori (Gaver 1993b). Han introduserer ideen om å bruke lyd i datamaskiner på en måte som gjenspeiler vår måte å gjenkjenne lydkilder i naturen på. Ideen går ut på at hvis forskere klarer å kartlegge hvilke fysiske attributter ved en gjenstand som påvirker hvordan en lyd høres ut, kan disse forskjellige attributtene danne en rekke med dimensjoner som lyddesignere kan bruke når de skal lage systemlyder. Dermed kan vi få lyder som kan beskrive objekters størrelse, masse, form, materiale og andre fysiske egenskaper. Gaver kaller denne typen systemlyder for *lydikoner*, fordi de på karikaturmessig vis hermer etter naturlige lyder ved å overdrive ulike dimensjoner i lyden (Gaver 1986).

Vi husker fra den multimodale analysen at den skjematiske sannhetsforsterkningen praktisk talt ikke var til stede i lydene. Gavers tilnærming vektlegger derimot nettopp slike attributter som kan omsettes til målbare størrelser, og at de overdrives slik at de karikerer de *materielle* egenskapene til objektene de skal forestille (Gaver 1989:75-76). Gavers syn er altså ikke så langt fra det vi kjenner som en skjematisk sannhetsforsterkning. Samtidig er Gavers tilnærming noe annerledes enn handling-lyd-slektskapet som jeg har fulgt gjennom denne oppgaven. Mens handling-lyd ser etter hva slags bevegelsesinntrykk som skapes av lyden, fokuserer Gaver på lydens tilknytning til ulike materialer, ikke ulikt det jeg i analysen kalte morfologiske kvaliteter. Gaver mener materialitet kan utnyttes i brukergrensesnittet ved å gi visse type objekter faste kvaliteter, for eksempel at dokumenter var laget av tre, programmer laget av metall, og mapper laget av papir (*ibid.*:78). Denne tilnærmingen kan fungere som et godt kompliment til handling-lyd-slektskapet i framtidig forskning, hvor den ene fokuserer

på bevegelse, mens den andre fokuserer på materialitet. Nedenfor vil jeg drøfte et eksempel fra Gavers forskning, for etterpå å sammenligne Gavers syn med funnene mine fra analysen.

### 5.3.2 SonicFinder

For å demonstrere en måte lydikoner kan implementeres i et brukergrensesnitt, har Gaver laget en modifisert utgave av filbehandleren Finder 6.0: “SonicFinder” (Gaver 1989). Jeg skal ta for meg et eksempel fra en handling derfra, nemlig det å flytte en fil fra et vindu og over til papirkurven.

Når en bruker klikker på filen, høres en impulsiv lyd av treverk som signaliserer at dette er et filobjekt. Lyden er lys eller mørk avhengig om filen er stor eller liten. Så begynner brukeren å dra i filen, og det høres en utholdt skrapelyd, hvor skrapelyden endrer karakter avhengig av om brukeren drar filen over et vindu eller over skrivebordet. Vinduer og skrivebord har altså en friksjonsparameter som påvirker frekvensinnholdet i lyden av det som dras over. Papirkurven er laget av metall, og det høres en hul metallisk impulsiv lyd når brukeren har dratt filen over papirkurven, etterfulgt av papirkrølling når brukeren slipper filen i papirkurven (ibid.:79-81).

Sammenlignet med det lydmiljøet vi er vant til fra dagens datamaskiner, kan ideen om å lydlegge alle objekter og handlinger i brukergrensesnittet virke skremmende. Hvordan kan man unngå lydmessig lapskaus? En viktig grunn til hvorfor jeg tror SonicFinder ikke ville fungert i dag, er at Finder 6.0 bare hadde mulighet til å kjøre et program om gangen.<sup>16</sup> Jeg vil hevde at etter at internett kom og dataskjermene ble større på 90-tallet, ble det vanlig for oss å ha flere programmer oppe samtidig, noe som med en mulig etterfølger til SonicFinder ville ført til et lydkaos som Gaver ikke kunne forutse ut fra datidens bruksmønstre og grensesnittparadigmer. Imidlertid kan teorien hans fortsatt være aktuell for dagens mobile enheter, fordi brukeren da tror han kjører ett program om gangen på grunn av begrensninger i skjermstørrelsen samt mindre presise betjeningsmuligheter enn for PCer.

Hvis vi likevel skulle ønske en karikert lydlegging av datamaskinens virtuelle verden i tråd med det Gaver foreslår, hvordan kan vi redusere det mulige lydkaoset som da kan oppstå i

---

<sup>16</sup> Preemptiv multitasking ble mulig i Macintosh omtrent på samme tid som Finder 6.0 kom ut, ved hjelp av et den gang ressurskrevende program som het MultiFinder

kombinasjonen mellom de ulike programmene? Jeg mener vi bør kikke litt over på hvordan tilsvarende problemer løses i brukergrensesnittets visuelle sfære. Der har man gjerne et aktivt vindu eller arbeidsområde som ligger i forgrunnen, og som maskerer de andre visuelle elementene som ligger i bakgrunnen. I lydverdenen finnes det også måter å maskere lyd på, blant annet maskering i tid, frekvens og filtrering av lyd. Direktivitet og ikke minst lydintensitet er også virkemidler som kan benyttes for å plassere lyder i forgrunnen eller bakgrunnen. Det bør altså undersøkes hva slags regler som skal styre hvilke lyder som brukeren jobber med, hvilke lyder som skal tilkalle brukerens oppmerksomhet og hvilke lyder som er mindre viktige og som danner et fjernt og svakt lydteppe. Dette er imidlertid en helt annen tankegang enn det som praktiseres av Microsoft i dag, og en sammenligning av motiver er på sin plass.

### 5.3.3 Sammenligning av Gaver og Microsoft

Når motivene til Microsoft og Gaver sammenlignes, er det to tydelige forskjeller som peker seg ut. For det første ser Microsoft lyd som et veldig oppmerksomhetskrevende virkemiddel. Hvis det brukes for mye lyd, overøses brukeren med forstyrrende elementer. Gaver ser på den annen side lyd som en økologisk faktor, hvor vi forstår lyd i brukergrensesnitt på samme måte som i den virkelige verden. Derfor vil vi ifølge Gaver ikke se en jevn strøm av lyder som forstyrrende, så lenge lydene står i samsvar til datamaskinens virtuelle verden. For det andre har Microsoft og Gaver svært forskjellige synspunkter på hvordan lyd skal få et brukergrensesnitt til å føles ut. Gaver ønsker at brukeren ikke skal se brukergrensesnittet, men heller føle at objektene manipuleres direkte. Microsoft ønsker derimot å synliggjøre brukergrensesnittet som et medieprodukt gjennom branding. Her er det altså en tydelig konflikt mellom lyd-usynliggjøring og lyd-branding, som ved første øyekast vil kunne by på en skikkelig utfordring for de som ønsker å designe systemlyder som blidgjør begge leire.

En løsning kan ligge i det [Jensenius \(2008\)](#) kaller praktiske og kreative designstrategier. Her påpekes det at en praktisk designstrategi fokuserer på det som er *enkelt* og lett å bruke, men blir lett for uinspirerende. En kreativ designstrategi fokuserer derimot på det som er *interessant* å bruke, men står i fare for å bli forvirrende. Disse designstrategiene danner to opposisjonelle akser, og det anbefales at lyddesignere bør søke en balansegang mellom disse to aksene ([Jensenius 2008:32](#)).

Ut fra funnene i denne oppgaven, tyder ting på at det ikke nødvendigvis er slik at praktiske og kreative designstrategier ekskluderer hverandre. Vi har sett eksempler på at lyder hverken er praktiske eller kreative, sånn som Notification-lyden til Windows Vista. Den lyden besto av en karakterløs lav lyd som det var vanskelig å utlede en ordentlig betydning ut av. Vi har også sett eksempler på at lyder både kan være praktiske og kreative, sånn som New Mail-lyden til Windows Vista. Den lyden var på sin side svært god til å formidle brukergrensesnittet Vista som et produkt ved å bruke kreative virkemidler som forholdt seg til Microsofts retningslinjer, samtidig som den var veldig praktisk takket være et fint bevegelsesinntrykk som eksisterte på tvers av flere samtidige modaliteter. Jeg mener at lyddesignere ikke bare må søke å balansere sin designstrategi mellom det praktiske og det kreative, men at det tas hensyn til å gjøre lyden praktisk *samtidig* som den blir gjort kreativ. Jeg så tendenser i den multimodale analysen på at forsterkning i en sannhetsgrad kunne smitte over på de andre, og på den måten styrke den generelle medlingsformidlingen. Tilsvarende tror jeg at praktisk designstrategi kan smitte over på kreativ designstrategi og vice versa, gitt at forskere og lyddesignere har en grunnleggende kjennskap til ulike virkemidler som kan brukes.

## 5.4 Kløften mellom forskning og programvareutvikling

Så langt i dette kapittelet har ting tydet på at det eksisterer en forskjell i motiver mellom hva forskere og hva programvareutviklere ser for seg at systemlyder kan brukes til. Det er klart at det er mulig å ønske seg det beste fra begge verdener, men spørsmålet er hvordan det kan være realiserbart.

Det virker som om både [Microsoft \(2008\)](#) og lærebøker i interaksjonsdesign ([Sharp et al. 2007](#)) ser på lyd som ‘prikken over i-en’ i brukergrensesnittet. Så snart et brukergrensesnitt er ferdig utviklet, er det bare å legge lydene på toppen av brukergrensesnittet for å gjøre brukergrensesnittet mer trivelig å bruke. Imidlertid vil en slik løsning virke hemmende for lydenes mulighet til å utvikle seg som effektiv meningsformidler, fordi lyden da er nødt til å følge det grafiske brukergrensesnittets regler. Hvis lyddesign i stedet inkluderes tidligere i utviklingsfasen, er det større mulighet for å motta tilbakemelding om de prinsippene som fungerer ifølge forskningen også fungerer i praksis. Da Windows 7 Public Beta ([Microsoft 2009](#)) ble gjort tilgjengelig for nedlasting tidlig i 2009, hadde det systemet akkurat de samme standardlydene som Windows Vista. En offentlig betaversjon ville vært en glimrende mulighet

for Microsoft til å få tilbakemelding fra brukerne om hvor godt systemlydene fungerte, men det ser ikke ut til at Microsoft benyttet seg av noen slik mulighet i dette tilfellet.

Problemet med å inkludere lyd tidlig i utviklingsfasen og offentliggjøre lyder i betaversjoner, er at det ikke er spesielt forenlig med en kommersiell utviklingsmodell. Et stort og kjent firma som Microsoft må forholde seg til flere ting enn bare å lage god programvare, det er også snakk om å møte produktspesifikasjoner, å holde løfter ovenfor interessenter og kunder, deadliner som må følges og PR som må opprettholdes for ikke å påvirke aksjekursen negativt. Forskningsmiljøer trenger normalt ikke å måtte bekymre seg for slike eksterne krefter, så her er det en forskjell mellom forskning og programvareutvikling som ikke har noen umiddelbar løsning. Det er mer sannsynlig at det er mulig for forskningsmiljøene å tilnærme seg programvareutviklere enn motsatt, da forskningen her står noe friere.

Hvordan kan forskere tilnærme seg programvareutviklere? Forskere kan ta hensyn til spillereglene som gjelder i kommersiell programvareutvikling, uten å la seg begrense av dem. Hvis det lages nok prototyper av programmer eller brukergrensesnitt med dristig lyddesign, vil det være en sjanse for at en av disse prototypene har noe som slår an. Det er ikke sikkert at virkemidlene i den mest populære prototypen nødvendigvis stemmer overens med det forskerne hadde sett for seg, men prototypen vil kunne sette økt fokus på å bruke lyddesign aktivt i brukergrensesnitt. Dermed vil det kunne redusere en eventuell frykt som programvareutviklere har mot å bruke for mye systemlyder, og mer praktisk-kreative løsninger kan dukke opp i kjølvannet av at programvareutviklerne i større grad inkluderer lyd tidlig i utviklingsprosessen.

Midt i alt fokuset på å inkludere lyd og visuelle inntrykk i brukergrensesnittet, er det viktig ikke å glemme at mennesket har fem sanser. En årsak til at lyd kanskje ikke har latt seg kombinere så godt med det grafiske brukergrensesnittet som forskere hadde ønsket, kan være at måten vi betjener datamaskiner på ikke egner seg godt til å gi lydmessig tilbakemelding. Et eksempel her kan komme fra mobiltelefoner med berøringsskjerm. Når brukeren skriver på et tastatur på en flat berøringsskjerm, mangler skjermen den taktile tilbakemeldingen som ble gitt av den fysiske motstanden mellom fingre og taster på et tastatur. Løsningen er å gi brukeren tilbakemelding via lyd, så det høres ut som om brukeren skriver på en tradisjonell skrivemaskin som lager en klikkelyd for hver bokstav. Et annet eksempel ligger i

bevegelsessensorer som vi finner i de samme mobiltelefonene. En bevegelsessensor registrerer endringer i posisjon, og det kan tolkes om til bevegelse. I denne oppgaven har vist hvordan slike bevegelsesinntrykk også kan vises i systemlyder. Hvis det derfor forskes på andre sanser enn hørsel og syn, kan dette virke stimulerende på forskningen på systemlyder. Jeg ser to årsaker til dette: for det første får brukeren flere måter å betjene utstyret på, og for det andre oppstår det mening i flere modaliteter, og lyden får dermed flere muligheter til å forankre sin betydning multimodalt.

Den siste måten som jeg ser til å få forskning og programvareutvikling nærmere hverandre, er ved å inkludere lydens meningsbærende egenskaper i undervisning som tar for seg interaksjonsdesign. Gitt denne oppgavens primære lesergruppe, er dette kanskje det viktigste argumentet mitt. Det er langt fra alle programvareutviklere som har profesjonell opplæring i interaksjonsdesign, men brukersentrert design har fått styrket sin posisjon i markedet de siste årene ([Sharp et al. 2007](#)). Hvis lærebokforfattere og lektorer anerkjenner at god lyddesign kan være lønnsomt å ha i brukergrensesnitt, og hvis det vises til forskning og litteratur på dette feltet, får morgendagens programvareutviklere med seg en viktig holdning fra forskningen. Som denne oppgaven har vist, finnes det allerede en god del teori og praksis som det kan refereres til.



## 6. Avslutning

---

Etter å ha kontekstualisert funnene mine i forrige kapittel, er det på tide oppsummere konklusjoner i denne oppgaven. Jeg har i løpet av oppgaven benyttet meg av mange teorier og metoder fra ulike fagtradisjoner, og jeg har brukt to problemstillinger, to analyser og en kontekstualisering for å komme fram til resultatene mine. En oppsummering av disse resultatene kan derfor komme godt med. Etter oppsummeringen vil jeg diskutere veien videre, ved kort å se på nærliggende forskning som pågår og komme med forslag på hvordan funnene i oppgaven kan brukes i senere forskning.

### 6.1 Oppsummering

Valget mitt om å kombinere metoder og begreper fra lydteori og semiotikk virket fornuftig, da jeg klarte å finne betydning i alle analyseobjektene mine. Likevel skal jeg ikke se bort fra at noen av resultatene kan ha blitt farget av at jeg kjente lydene og hendelsene som jeg analyserte, og at jeg derfor kanskje lette bevisst etter enkelte betydninger. Når jeg ser hvor sprikende resultatene ble, ville jeg nok ikke klart å se noen sammenheng mellom funnene mine hvis jeg hadde analysert lydene uten å kjenne til hvilke situasjoner de tilhørte. Uansett tror jeg at det ble skapt en god analytisk balanse ved at jeg studerte lydene fra tre forskjellige perspektiver: lydens betydning i seg selv, lydens betydning i dens sammenheng mot andre medierte inntrykk, for til slutt å kontekstualisere funnene.

Den første problemstillingen, hvordan semiotiske trekk i Windows har utviklet seg de siste årene, mener jeg at jeg lyktes med å besvare, til tross for at resultatene pekte i mange retninger. I Windows 2000 var det bare noen få lyder, og dette resulterte i at noen lyder hadde samme betydning for hendelser som var vidt forskjellige. I Windows XP var det en stor variasjon i lydenes uttrykksform. Bevegelsene var tydeligere enn i de andre systemene, og det var stor variasjon i lydstyrke og i lydobjektens ulike kvaliteter. I Windows Vista var lydene mer lik hverandre igjen, fordi lydstyrken ble jevnere, anslaget mykere og bassinnholdet redusert i alle lydene i forhold til tidligere systemer. Dette ble gjort selv om forandringen gikk på bekostning av meningsformidlingen, hvor lyder som var enkle å forstå i Windows XP ble vanskeligere å forstå i Windows Vista. Det som derimot ble lettere å forstå i Windows Vista, var at likheten

mellom lydene fikk fram en spesiell brukeropplevelse som var ment å gjøre brukeren oppmerksom på medieproduktet Windows Vista. Dette er ikke et valg som er særlig lett å kombinere med idealer om brukervennlighet, men det er fornuftige valg ut fra dagens kommersielle utviklersituasjon. Likevel fant jeg tegn på at de intervallbaserte systemlydene var velegnet til *både* å inneholde betydning om handlinger i brukergrensesnittet og betydning om den medierte brukeropplevelsen. Disse lydene indikerer etter min mening at det finnes enkelte konvensjoner for systemlyd som holdes ved like på tvers av operativsystemene, men jeg fant flere eksempler på at det ikke tas hensyn til betydning i tidligere systemlyder når nye systemlyder skal lages for et nytt operativsystem.

Problemstilling nummer to, hva som kjennetegner systemlydenes betydningsstrøm, ga noe mer uklare resultater. Det var lite hensiktsmessig for meg å se etter konkrete lineære betydningsstrømmer, da disse viste seg å være svært forskjellige fra det ene analyseobjektet til det neste. Det var også vanskelig å finne at lyden hadde noen konnoterende betydning, og en mulig årsak til dette fant jeg i kontekstualiseringen da jeg påpekte at Microsoft ønsker at systemlyder helst ikke skal konnotere, i frykt for at lydene skal mislikes av enkelte brukere. Når det gjaldt systemlydenes sannhetsforsterkning, gjorde jeg noen etter min mening interessante funn. En høy naturalistisk sannhetsforsterkning gjorde det mulig å finne betydning i lyden som ble overført til skrift og bilde, mens en høy abstrakt sannhetsforsterkning førte til at betydningen ble overført motsatt vei, fra skrift og bilde til lyd. Her er det nødvendig å påpeke at jeg ikke hadde tydelige nok resultater i analyseobjektene mine, og funnene mine om sannhetsforsterkning vil måtte trenge flere bekreftende analyser hvis de skal stå sterkere.

Det er alltid en fare ved å kombinere teori fra flere ulike forskningstradisjoner, hvor begreper som ligner hverandre ikke nødvendigvis er kompatible. Et eksempel på et slikt blandingsbegrep er 'bevegelse', hvor begrepet spiller en rolle både i lydteori (handling-lyd-slektskap og lydobjekter i bevegelse), i semiotikk (mening som beveger seg multimodalt) og i metaforteleori (romlighetsmetaforer og *image schemata*). Jeg mener at mitt valg om å se etter bevegelsesinntrykk i analysen ikke kom i konflikt med de ulike synene på hva bevegelse kan være, selv om det er mulig at dette valget har skapt litt mindre presisjon i argumentasjonen. Likevel lyktes jeg med å vise at de lydene som konserverte et bevegelsesinntrykk på tvers av flere modaliteter, var de lydene som var lettest å forstå og var også de lydene som hadde minst radikale endringer fra et system til et annet. Derfor tror jeg at mitt valg om å bruke forenklete

blandingsbegreper i argumentasjonen resulterte i et bedre overblikk og en bedre mulighet til å se resultatene mine i en større sammenheng.

## 6.2 Veien videre

Som jeg nevnte i **Kapittel 1**, begynner forskning på menneskets lydforståelse å skyte fart i ulike forskningsmiljøer, og denne oppgaven tar del i denne utviklingen. Under arbeidet med masteroppgaven har jeg kommet over et nærliggende forskningsprosjekt som er verdt å nevne. Prosjektet “Sensing Music-related Actions” (<http://www.hf.uio.no/imv/forskning/forskningsprosjekter/musicalactions/>) er et samarbeid ved Universitetet i Oslo mellom Institutt for musikkvitenskap og Institutt for informatikk, og har pågått siden sommeren 2008. Selv om Sensing Music-related Actions har mer fokus på musikk og teknologi enn på systemlyder i brukergrensesnitt, baserer prosjektet seg på den samme forståelsen av handling-lyd-slektskapet som jeg la til grunn for denne oppgaven. Det er mulig at noen av resultatene i dette prosjektet kan anvendes til videre forskning på denne oppgavens problemområde, for eksempel å bruke tekniske løsninger for å koble bevegelser og musikk sammen i et brukergrensesnitt. Det finner også sted en del årlige konferanser innenfor relaterte områder, deriblant “International Community for Auditory Display” (<http://www.icad.org>) og “Quality of Multimedia Experience” (<http://www.qomex.org>). Alt i alt tyder aktiviteten innenfor disse forskningsfeltene på at lyd er langt mer anerkjent som meningsbærer nå enn for bare noen få år siden.

Det er et par områder som denne oppgaven tok opp, hvor funnene forhåpentlig kan gi inspirasjon til videre forskning. Den første innfallsvinkelen jeg vil foreslå, er å prøve sammensetningen av handling-lyd-slektskapet og semiotikk på noen andre analyseobjekter. Enten kan analyseobjektene utvides til også å studere de musikalske påloggingslydene som jeg valgte å se bort fra, for å se om det har vært en lignende historisk utvikling der. Et annet analyseobjekt kan være brukergrensesnittet i mobiltelefoner, et eksempel som jeg nevnte i kontekstualiseringen min. En tredje mulighet er å se tilbake i tid, til hvordan systemlyder og lydeffekter fungerte i spillkonsollenes barndom. Noen av dagens teknisk krevende løsninger i lyddesign var ikke anvendbare den gang, men, som jeg påpekte i **Kapittel 5.2.1**, var det fullt mulig å lage intervallbaserte lyder med datidens teknologi. En slik analyse kan dessuten kombineres med å analysere lydene i en moderne spillkonsoll, for å se historisk utvikling i et lengre perspektiv.

Den andre innfallsvinkelen jeg vil foreslå er å bruke de samme analyseobjektene som jeg har brukt, men å benytte seg av andre metoder. Et eksempel kan være å inkludere funnene mine i en oppgave som tar for seg hvordan Microsoft bruker branding som en markedsføringsstrategi, og konsekvensene dette får for brukeropplevelsen. En annen mulighet er å se systemlydene i en mer sosiosemiotisk sammenheng. Da går det for eksempel an å studere hvordan muligheten til å personalisere lyder i Windows, eller kanskje i enda større grad mobiltelefoner, påvirker meningsformidlingen, eventuelt knyttet opp mot at datamaskinen/mobiltelefonen blir et uttrykk for brukerens identitet. Til slutt er det selvfølgelig mulig å teste funnene mine ved å produsere egne prototyper på brukergrensesnitt og systemlyder som benytter seg av lydegenskaper som jeg lette etter i denne oppgaven. Dette vil kunne være realiserbart enten på en nettside eller i et program til mobiltelefoner med berøringsskjerm hvor brukerens berøring av skjermen vil kunne tilføre enda et bevegelsesinntrykk som kan kombineres med bevegelsesinntrykket som ligger i lydene.

I denne oppgaven har jeg villet vise hvordan betydning kan ligge selv i lyder som er svært korte og lite oppmerksomhetskrevenende. Mye av denne betydningen kan spores tilbake til et bevegelsesinntrykk som sier noe om en handling som kan ha skapt lyden, enten i brukerverdenen eller i den virtuelle verdenen i brukergrensesnittet. Jeg håper metodene som jeg brukte til å finne semiotisk betydning i lydene vil være til hjelp for andre som skal foreta lignende analyser. Mitt analyseområde var systemlyder, men dette er selvfølgelig bare en liten kategori av lyder i det totale mediebildet. Jeg har vist at lyden spiller en viktig rolle i meningsformidlingen når den inntreffer sammen med skrift og bilde i et brukergrensesnitt, og hvis tilsvarende regler gjelder i andre medier som internett og film, mener jeg vi kan slå fast at lyd er en meningsbærer som det hørende menneske kan nyttiggjøre i sin hverdag.

# Referanser

---

- Barthes, Roland (1977a) "The Photographic Message" i *Image, Music, Text* London: Fontana  
Originaltittel: "Le message photographique" 13
- Barthes, Roland (1977b) "The Rhetoric of the Image" i *Image, Music, Text* London: Fontana  
Originaltittel: "Rhétorique de l'image" 13-15
- Barthes, Roland (1977c) "The Grain of the Voice" i *Image, Music, Text* London: Fontana  
Originaltittel: "Le grain de la voix" 2
- Barthes, Roland ([1957] 2002) *Mytologier* [Originaltittel: Mythologies] Oslo: Gyldendal Norsk Forlag 12-13
- Boersma og Weenink (2009) "PRAAT 5.1" [Programvare] University of Amsterdam: Institute of Phonetic Sciences. Tilgjengelig: <http://www.fon.hum.uva.nl/praat/> [2009, 1. mai] 22
- Chion, Michel ([1983] 1995) *Guide des objets sonores* Paris: Buchet/Chastel 8
- Gaver, William W. (1986) "Auditory Icons: Using Sound in Computer Interfaces" *Human-Computer Interaction*, 2(2), 167-177 79
- Gaver, William W. (1989) "The SonicFinder: An Interface That Uses Auditory Icons" *Human-Computer Interaction*, 4(1), 67-94 79-80
- Gaver, William W. (1993a) "How Do We Hear in the World? Explorations in Ecological Acoustics" *Ecological Perception*, 5(4), 285-313 28
- Gaver, William W. (1993b) "What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception" *Ecological Perception*, 5(1), 1-29 79
- Gibson, J. J. (1979) *The Ecological Approach to Visual Perception* New York: Houghton-Mifflin 7
- Giordano, Bruno L. (2003) "Material categorization and hardness scaling in real and synthetic impact sounds" i (Red) Rocchesso, Davide og Fontana, Federico *The Sounding Object* Florence: Edizioni di Mondo Estremo ss. 73-93 31
- Godøy, Rolf Inge (1997) *Formalization and Epistemology* Oslo: Universitetsforlaget 50
- Godøy, Rolf Inge (2006) "Sonorous Objects: Embodied extensions of Schaeffer's conceptual apparatus" i *Organized Sound* 2006 volum 11(2) ss. 149-157 8-11, 19, 23, 31
- Grassi, Massimo og Burro, Roberto (2003) "Impact Sounds" i (Red) Rocchesso, Davide og Fontana, Federico *The Sounding Object* Florence: Edizioni di Mondo Estremo ss. 47-69 23, 40

- Hitchcock, Alfred (1960) *Psycho* [Film] (USA) 18
- Hjelmevold, Bjørnar (2009) “Lydmatrise” [Online] Tilgjengelig: <http://folk.uio.no/bjorla/master/> [2009, 1. mai] Alternativ: <http://homepage.mac.com/bjornarlassen/master/> 29
- Hull, Glynda A. og Nelson, Mark Evan (2005) “Locating the Semiotic Power of Multimodality” i *Written Communication* California: Sage Publications 3-4
- Jensenius, Alexander Refsum (2008) *Action – Sound: Devolping Methods and Tools to Study Music-related Body Movement* PhD-avhandling. Oslo: UiO, Institutt for musikkvitenskap 8-9, 11, 28, 81
- Johnson, Mark (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* Chicago: University of Chicago Press 18-19, 44-45
- Kress, Günther og Van Leeuwen, Theo (2001) *Multimodal Discourse. The modes and media of contemporary communication* New York: Oxford University Press 4, 16-18, 36, 38
- Lakoff, George og Johnson, Mark ([1980] 2003) *Hverdagslivets metaforer. Fornuft, følelser, og menneskehjernen* [Originaltittel: *Metaphors We Live By*] Oslo: Pax Forlag 18-19, 32, 36
- Larsen, Peter (1988) “Betydningsstrømme. Musik og moderne billedfiksjoner” i *Studia Musicologica Norvegica. Norsk årsskrift for musikkforskning* (14) 1988 ss. 19-52 14-15, 61
- Leman, Mark (2007) *Embodied Music Cognition and Mediation Technology* Cambridge, MA: MIT Press 56
- Lending, Mari (2002) “Innledende Essay” i Barthes, Roland ([1957] 2002) *Mytologier* Oslo: Gyldendal Norsk Forlag 12-13
- Microsoft Corporation (1999) “Windows 2000 Professional Service Pack 4 (English)” [Programvare] Redmond, WA, USA 4
- Microsoft Corporation (2001) “Windows XP Professional Service Pack 3 (English)” [Programvare] Redmond, WA, USA 4, 58, 62, 68
- Microsoft Corporation (2006) “Windows Vista Home Premium Service Pack 1 (English)” [Programvare] Redmond, WA, USA 4, 66
- Microsoft Corporation (2008) *Windows Vista User Experience Guidelines* [Online PDF] Tilgjengelig: <http://msdn.microsoft.com/en-us/library/aa511258.aspx> [2009, 1. mai] 1-2, 23-24, 71-77
- Microsoft Corporation (2009) “Windows 7 Public Beta Version 7.000 64-bit (English)” [Programvare] Redmond, WA, USA 82
- Maasø, Arnt (1995) *Lyden av levende bilder* IMK-rapport nr. 14, Institutt for medier og kommunikasjon. Oslo: UiO 3, 15, 75

- Maasø, Arnt (2002) “This goes to eleven” i (Red) Ericson, S. og Ytreberg, E. *Fjernsyn mellom høy og lav kultur* Oslo: Høyskoleforlaget ss. 175-204 51
- Nattiez, Jean-Jacques (1990) *Music and Discourse: Toward a Semiology of Music* New Jersey: Princeton University Press 8
- Norman, Donald ([1988] 2002) *The Design of Everyday Things* New York: Basic Books 7, 78-79
- Norman, Donald (2004) *Emotional Design* New York: Basic Books 73
- Reimer, Jeremy (2006) “Gentlemen, start your Vista chimes (or stop them)” [Online] Tilgjengelig: <http://arstechnica.com/business/news/2006/09/7815.ars> [2009, 1. mai] 73
- Rocchesso, Davide og Fontana, Federico (2003) *The Sounding Object* Florence: Edizioni di Mondo Estremo 2, 78
- Schaeffer, Pierre (1966) *Traité des objets musicaux* Paris: Éditions du Seuil 7, 9-10
- Sharp, Helen; Rogers, Yvonne; Preece, Jeniffer (2007) *Interaction Design: Beyond human-computer interaction*, 2nd edition. Hoboken, NJ: Wiley & Sons 2, 64, 82, 84
- Snyder, Bob (2000) *Music and Memory. An Introduction*. London: MIT Press 19, 75
- Storey, John (2006) *Cultural Theory and Popular Culture: An Introduction* Harlow: Pearson Education 7, 11-12
- Svanæs, Dag (2000) *Understanding Interactivity. Steps to a Phenomenology of Human-Computer Interaction*. Avhandling (dr. philos.) Trondheim: Norges teknisk-naturvitenskapelige universitet 7
- Tagg, Philip (2008) “Intel Inside Analysis (Full Version)” [Online] Tilgjengelig: <http://www.tagg.org/ptavmat.htm#IntelNew> [2009, 1. mai] 29
- Torre, Charles (2006) “Behind the scenes at Windows Vista recording session” [Online Video] Tilgjengelig: <http://channel9.msdn.com/posts/scobleizer/Robert-Fripp-Behind-the-scenes-at-Windows-Vista-recording-session/> [2009, 1. mai] 73
- Van Leeuwen, Theo (1999) *Speech, Music, Sound* London: McMillan 16, 51, 54, 67